



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification: C12Q		A2	(11) International Publication Number: WO 00/75358 (43) International Publication Date: 14 December 2000 (14.12.2000)
(21) International Application Number: PCT/US00/15364			
(22) International Filing Date: 02 June 2000 (02.06.2000)		Published	
(30) Priority Data: 60/137,659 04 June 1999 (04.06.1999) US			
(60) Parent Application or Grant MILLENNIUM PHARMACEUTICALS, INC. [/]; O. KHODADOUST, Mehran [/]; O. MANDRAGOURAS, Amy, E. ; O.			
<p>(54) Title: NOVEL TOLL MOLECULES AND USES THEREFOR (54) Titre: MOLECULES TOLL ET LEURS UTILISATIONS</p> <p>(57) Abstract</p> <p>The invention provides isolated nucleic acid molecules, designated TOLL nucleic acid molecules, which encode novel TOLL family members. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing TOLL nucleic acid molecules, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a TOLL gene has been introduced or disrupted. The invention still further provides isolated TOLL proteins, fusion proteins, antigenic peptides and anti-TOLL antibodies. Diagnostic methods utilizing compositions of the invention are also provided.</p> <p>(57) Abrégé</p> <p>Cette invention concerne des molécules d'acides nucléiques isolées dites molécules d'acides nucléiques TOLL, qui codent pour des membres de la famille TOLL. L'invention porte également sur des molécules d'acides nucléiques antisens, des vecteurs d'expression recombinants renfermant des molécules d'acides nucléiques TOLL, des cellules hôtes dans lesquels on a introduit des vecteurs d'expression et des animaux transgéniques non humains dans lesquels un gène TOLL a été introduit ou disloqué. Un autre aspect de l'invention concerne des protéines TOLL isolées, des protéines hybrides, des peptides antigéniques et des anticorps anti-TOLL. Sont également décrites des méthodes faisant intervenir des compositions selon l'invention.</p>			

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 December 2000 (14.12.2000)

PCT

(10) International Publication Number
WO 00/75358 A2

(51) International Patent Classification⁷:

C12Q

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(21) International Application Number:

PCT/US00/15364

(22) International Filing Date:

2 June 2000 (02.06.2000)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/137,659

4 June 1999 (04.06.1999) US

(71) Applicant: MILLENNIUM PHARMACEUTICALS, INC. [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US).

(72) Inventor: KHODADOUST, Mehran; 106 Sewall Avenue #6, Brookline, MA 02445 (US).

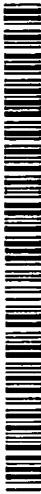
(74) Agents: MANDRAGOURAS, Amy, E.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 et al. (US).

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 00/75358 A2

(54) Title: NOVEL TOLL MOLECULES AND USES THEREFOR

(57) Abstract: The invention provides isolated nucleic acid molecules, designated TOLL nucleic acid molecules, which encode novel TOLL family members. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing TOLL nucleic acid molecules, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a TOLL gene has been introduced or disrupted. The invention still further provides isolated TOLL proteins, fusion proteins, antigenic peptides and anti-TOLL antibodies. Diagnostic methods utilizing compositions of the invention are also provided.

Description

5

10

15

20

25

30

35

40

45

50

55

5

- 1 -

NOVEL TOLL MOLECULES AND USES THEREFOR

10

Background of the Invention

The Toll protein in *Drosophila* is an integral membrane receptor involved in many different cellular processes ranging from cellular adhesion (Keith, F.J. et al., (1990) *EMBO J.* 9:4299-4306) to muscle formation (Halfon, M.S.. et al., (1998) *Dev. Biol.* 199:164-174) to dorsal-ventral patterning during development (Hashimoto, C. et al., (1988) *Cell* 52:269-279). Studies in various other organisms have identified a number of related proteins, including the *Drosophila* 18-wheeler (Eldon, E., et al., (1994) *Development* 120:885-899), MstProx (Hashimoto, C. et al., (1988) *Cell* 52:269-279), and STSDm2245 proteins (Mitcham, J.L., et al., (1996) *J. Biol. Chem.* 271:5777-5783), the tobacco N gene product (Whitham, S. et al., (1994) *Cell* 78:1101-1115), mammalian and avian type-1 interleukin-1 receptors (Yamagata, M., et al. (1994) *Gene* 139:223-8), and human Toll-like proteins 1-5 (Rock, F.L., et al., (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:588-593).

30

Family members contain one or more specific conserved domains found in Toll. These include an extracellular leucine-rich repeat region, a domain containing a β/α -class fold with α -helices on both faces of a parallel β sheet, and a transmembrane domain (Hashimoto, C. et al., (1988) *Cell* 52:269-279; Rock, F.L., et al., (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:588-593). The specific functions of each of these domains within Toll have been explored. There is evidence that the extracellular leucine rich repeat region may promote cell to cell contact and adhesion due to the tendency of this domain to form an amphipathic structure with a predominantly apolar and charged surface (Gay, N.J., et al., (1991) *FEBS Lett* 291:87-91). The transmembrane domain serves to anchor the receptor into the membrane, and the cytosolic β/α -class fold domain is posited to be a conformational trigger for signaling (Rock, F.L., et al., (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:588-593).

35

The conserved cytoplasmic domain between Toll and the avian/mammalian type 1 IL-1 receptor, containing the β/α -class fold, suggested another potential function for 40 Toll. The IL-1 receptor is involved in activation of the transcription factor NF κ B in 45

50

55

5 response to stimulation by IL-1 in an innate immune response (O'Neill, L.A. et al.,
(1998) *J. Leukoc. Biol.* 63: 650-657). It was found that mutation of certain residues in
10 this domain not only abrogated the signaling activity of the type 1 IL-1 receptor, but also
similarly impaired Toll function (Heguy, A. et al., (1992) *J. Biol. Chem.* 267:2606-
15 2609). Further studies determined that Toll controls an immune response to fungal
infection in *Drosophila* (Lemaitre, B., et al., (1996) *Cell* 86:973-983) and that 18-
wheeler is critical for the normal functioning of the antibacterial response in these
20 organisms (Williams, M.J., et al., (1997) *EMBO J.* 16:6120-6130). Similarly, studies of
the tobacco N protein have demonstrated that this protein mediates resistance to tobacco
25 mosaic virus (Whitham, S. et al., (1994) *Cell* 78:1101-1115). Thus, Toll family
members may play key roles in host resistance to infection or in the activation of the
immune response in general.

25 **Summary of the Invention**

15 The present invention is based, at least in part, on the discovery of novel TOLL
family members, referred to herein as "TOLL" nucleic acid and protein molecules. The
30 TOLL molecules of the present invention are useful as targets for developing
modulating agents to regulate a variety of cellular processes. Accordingly, in one
aspect, this invention provides isolated nucleic acid molecules encoding TOLL proteins
20 or biologically active portions thereof, as well as nucleic acid fragments suitable as
35 primers or hybridization probes for the detection of TOLL-encoding nucleic acids.

In one embodiment, a TOLL nucleic acid molecule of the invention is at least
50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or more identical to the
40 nucleotide sequence (e.g., to the entire length of the nucleotide sequence) shown in SEQ
25 ID NO:1, 3, or 4 or the nucleotide sequence of the DNA insert of the plasmid deposited
with ATCC as Accession Number _____, or a complement thereof.

In a preferred embodiment, the isolated nucleic acid molecule includes the
45 nucleotide sequence shown SEQ ID NO:1, 3, or 4, or a complement thereof. In another
embodiment, the nucleic acid molecule includes SEQ ID NO:3 and nucleotides 1-91 of
30 SEQ ID NO:1. In another embodiment, the nucleic acid molecule includes SEQ ID
NO:3 and nucleotides 1736-2147 of SEQ ID NO:1. In another preferred embodiment,
50

5

the nucleic acid molecule consists of the nucleotide sequence shown in SEQ ID NO:1, 3, or 4. In another preferred embodiment, the nucleic acid molecule includes a fragment of at least 491 nucleotides (e.g., 491 contiguous nucleotides) of the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof.

10

5 In another embodiment, a TOLL nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence sufficiently homologous to the amino acid sequence of SEQ ID NO:2 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In a preferred embodiment, a TOLL nucleic acid molecule includes a nucleotide sequence 15 encoding a protein having an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to the entire length of the amino acid sequence of SEQ ID NO:2 or the amino acid sequence encoded by the DNA insert 20 of the plasmid deposited with ATCC as Accession Number _____.
25

20

In another preferred embodiment, an isolated nucleic acid molecule encodes the 15 amino acid sequence of human TOLL. In yet another preferred embodiment, the nucleic acid molecule includes a nucleotide sequence encoding a protein having the amino acid sequence of SEQ ID NO:2 or the amino acid sequence encoded by the DNA insert of the 30 plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, the nucleic acid molecule is at least 491 nucleotides in length. In a further 20 preferred embodiment, the nucleic acid molecule is at least 491 nucleotides in length and 35 encodes a protein having a TOLL activity (as described herein).

30

Another embodiment of the invention features nucleic acid molecules, preferably 40 TOLL nucleic acid molecules, which specifically detect TOLL nucleic acid molecules relative to nucleic acid molecules encoding non-TOLL proteins. For example, in one 25 embodiment, such a nucleic acid molecule is at least 491, 491-500, 500-600, 600-700, 700-800, 800-900, 900-1000, 1000-1100, 1100-1200, 1200-1300, 1300-1400, or 1400-45 1500 or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence shown in SEQ ID NO:1, the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as 30 Accession Number _____. or a complement thereof.

50

55

5

In preferred embodiments, the nucleic acid molecules are at least 15 (e.g., contiguous) nucleotides in length and hybridize under stringent conditions to nucleotides 1-547, or 1039-1322 of SEQ ID NO:1, or to nucleotides 1-763 of SEQ ID NO: 4. In other preferred embodiments, the nucleic acid molecules comprise nucleotides 1-547, or 5 1039-1322 of SEQ ID NO:1, or nucleotides 1-763 of SEQ ID NO: 4.

10

In other preferred embodiments, the nucleic acid molecule encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the nucleic acid molecule hybridizes 10 to a nucleic acid molecule comprising SEQ ID NO:1, 3, or 4 under stringent conditions.

20

Another embodiment of the invention provides an isolated nucleic acid molecule which is antisense to a TOLL nucleic acid molecule, e.g., the coding strand of a TOLL nucleic acid molecule.

25

Another aspect of the invention provides a vector comprising a TOLL nucleic acid molecule. In certain embodiments, the vector is a recombinant expression vector. In another embodiment, the invention provides a host cell containing a vector of the invention. In yet another embodiment, the invention provides a host cell containing a nucleic acid molecule of the invention. The invention also provides a method for 30 producing a protein, preferably a TOLL protein, by culturing in a suitable medium, a host cell, e.g., a mammalian host cell such as a non-human mammalian cell, of the invention containing a recombinant expression vector, such that the protein is produced.

35

Another aspect of this invention features isolated or recombinant TOLL proteins and polypeptides. In one embodiment, the isolated protein, preferably a TOLL protein, includes at least one transmembrane domain. In another embodiment, the isolated 40 protein, preferably a TOLL protein, includes a leucine-rich repeat (LRR) domain. In yet another embodiment, the isolated protein, preferably a TOLL protein, includes a domain containing a β/α -class fold with α -helices on both faces of a parallel β sheet. In a preferred embodiment, the protein, preferably a TOLL protein, includes at least one transmembrane domain and has an amino acid sequence at least about 50%, 55%, 60%, 45 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to the amino acid 30 sequence of SEQ ID NO:2 or the amino acid sequence encoded by the DNA insert of the 50

5 plasmid deposited with ATCC as Accession Number _____. In another preferred embodiment, the protein, preferably a TOLL protein, includes a leucine-rich repeat (LRR) domain and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to the amino acid sequence

10 5 of SEQ ID NO:2 or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another preferred embodiment, the protein, preferably a TOLL protein, includes a domain containing a β/α -class fold with α -helices on both faces of a parallel β sheet and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more

15 10 homologous to the amino acid sequence of SEQ ID NO:2 or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, the protein, preferably a TOLL protein, includes at least one transmembrane domain, at least one leucine-rich repeat (LRR) domain and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to the amino acid sequence of SEQ ID

20 15 NO:2 or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, the protein, preferably a TOLL protein, includes at least one transmembrane domain and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes

25 30 20 under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, or 4. In a further embodiment, the protein, preferably a TOLL protein, includes a leucine-rich repeat domain and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence

35 40 25 of SEQ ID NO:1, 3, or 4. In yet another preferred embodiment, the protein, preferably a TOLL protein, includes at least one domain containing a β/α -class fold with α -helices on both faces of a parallel β sheet and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1, 3, or 4. In

45 30 50 yet another preferred embodiment, the protein, preferably a TOLL protein, includes at

5

least one transmembrane domain, at least one leucine-rich repeat (LRR) domain and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1, 3, or 4.

10

5 In another embodiment, the invention features fragments of the protein having the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 15 amino acids (e.g., contiguous amino acids) of the amino acid sequence of SEQ ID NO:2 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number _____. In another embodiment, the protein, preferably a 10 TOLL protein, has the amino acid sequence of SEQ ID NO:2, respectively.

20

In another embodiment, the invention features an isolated protein, preferably a TOLL protein, which is encoded by a nucleic acid molecule consisting of a nucleotide sequence at least about 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 25 more homologous to a nucleotide sequence of SEQ ID NO:1,3 or 4, or a complement 15 thereof. This invention further features an isolated protein, preferably a TOLL protein, which is encoded by a nucleic acid molecule consisting of a nucleotide sequence which 30 hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, or 4, or a complement thereof.

30

The proteins of the present invention or portions thereof, e.g., biologically active 35 portions thereof, can be operatively linked to a non-TOLL polypeptide (e.g., heterologous amino acid sequences) to form fusion proteins. The invention further features antibodies, such as monoclonal or polyclonal antibodies, that specifically bind 40 proteins of the invention, preferably TOLL proteins. In addition, the TOLL proteins or biologically active portions thereof can be incorporated into pharmaceutical 45 compositions, which optionally include pharmaceutically acceptable carriers.

40

In another aspect, the present invention provides a method for detecting the 45 presence of a TOLL nucleic acid molecule, protein or polypeptide in a biological sample by contacting the biological sample with an agent capable of detecting a TOLL nucleic acid molecule, protein or polypeptide such that the presence of a TOLL nucleic acid 50 molecule, protein or polypeptide is detected in the biological sample.

50

55

5

In another aspect, the present invention provides a method for detecting the presence of TOLL activity in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of TOLL activity such that the presence of TOLL activity is detected in the biological sample.

10

5 In another aspect, the invention provides a method for modulating TOLL activity comprising contacting a cell capable of expressing TOLL with an agent that modulates TOLL activity such that TOLL activity in the cell is modulated. In one embodiment, the agent inhibits TOLL activity. In another embodiment, the agent stimulates TOLL activity. In one embodiment, the agent is an antibody that specifically binds to a TOLL 15 protein. In another embodiment, the agent modulates expression of TOLL by modulating transcription of a TOLL gene or translation of a TOLL mRNA. In yet another embodiment, the agent is a nucleic acid molecule having a nucleotide sequence 20 that is antisense to the coding strand of a TOLL mRNA or a TOLL gene.

20

25 In one embodiment, the methods of the present invention are used to treat a subject having a disorder characterized by aberrant or unwanted TOLL protein or nucleic acid expression or activity by administering an agent which is a TOLL modulator to the subject. In one embodiment, the TOLL modulator is a TOLL protein. In another embodiment the TOLL modulator is a TOLL nucleic acid molecule. In yet another embodiment, the TOLL modulator is a peptide, peptidomimetic, or other small 30 molecule.

30

35 The present invention also provides a diagnostic assay for identifying the presence or absence of a genetic alteration characterized by at least one of (i) aberrant modification or mutation of a gene encoding a TOLL protein; (ii) mis-regulation of the gene; and (iii) aberrant post-translational modification of a TOLL protein, wherein a 40 wild-type form of the gene encodes a protein with a TOLL activity.

40

45 In another aspect the invention provides a method for identifying a compound that binds to or modulates the activity of a TOLL protein, by providing an indicator composition comprising a TOLL protein having TOLL activity, contacting the indicator composition with a test compound, and determining the effect of the test compound on 50 30 TOLL activity in the indicator composition to identify a compound that modulates the activity of a TOLL protein.

55

5

Other features and advantages of the invention will be apparent from the following detailed description and claims.

10

5 Brief Description of the Drawings

15

Figure 1 depicts the cDNA sequence and predicted amino acid sequence of human TOLL. The nucleotide sequence corresponds to nucleic acids 1 to 2147 of SEQ ID NO:1. The amino acid sequence corresponds to amino acids 1 to 548 of SEQ ID NO:2. The coding region without the 5' and 3' untranslated regions of the human TOLL

20

10 gene is shown in SEQ ID NO:3.

25

Figure 2 depicts the results of a search which was performed against the HMM database and which resulted in the identification of a leucine-rich repeat domain in the human TOLL protein.

30

15 *Figure 3* depicts the nucleotide sequence of mouse TOLL. The nucleotide sequence corresponds to nucleic acids 1 to 763 of SEQ ID NO:4.

30 Detailed Description of the Invention

35

The present invention is based, at least in part, on the discovery of novel nucleic acid and protein molecules which may be Toll family members, referred to herein as

40

20 "TOLL" nucleic acid and protein molecules. Due to their homology with Toll family members, the TOLL molecules of the present invention may be membrane proteins, e.g., membrane proteins which function as receptors, and they may be involved in immune signaling mechanisms.

45

25 As used herein, the term "immune signaling mechanisms" includes the cellular mechanisms involved in the development and regulation of an immune response, e.g., an immune response to the presence of a pathogen or antigen in a subject, e.g., a mammal such as a human. In mammals, the initial detection of a "foreign" pathogen or antigen results in the triggering of the innate immune response, in which specialized cells are recruited to engulf and destroy the invader, while chemical signals are simultaneously 30 generated and released to stimulate the activation of the acquired immune system. The TOLL molecules of the present invention may be involved in pathways involved in the

50

55

5

initial detection of these invaders, and/or in the recruitment of specialized cells such as macrophages, and/or in stimulation of cytokine production and release.

10

Thus, the TOLL molecules, by participating in immune signaling mechanisms, may be useful in the development of novel diagnostic targets and therapeutic agents to regulate immune and inflammatory responses in a variety of disorders, diseases, or conditions which are characterized by a deregulated, e.g., upregulated or downregulated, immune or inflammatory response. For example, the TOLL molecules may provide novel diagnostic targets and therapeutic agents for controlling disorders, diseases, or conditions related to misregulation of an immune response, such as rheumatoid arthritis, systemic lupus erythematosus, myasthenia gravis, Grave's disease, Sjogren syndrome, polymyositis and dermatomyositis, psoriasis, pemphigus vulgaris, bullous pemphigoid, inflammatory bowel disease, Kawasaki disease, asthma, and graft vs. host disease.

15

The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. Such family members can be naturally or non-naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of human origin or alternatively, can contain homologues of non-human origin. Members of a family may also have common functional characteristics.

20

For example, the family of TOLL proteins include at least one "transmembrane domain". As used herein, the term "transmembrane domain" includes an amino acid sequence of about 15 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an α -helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W.N. et al, (1996) *Annual Rev. Neurosci.* 19: 235-63, the contents of which are incorporated

25

30

35

5

herein by reference. Amino acid residues 456-480 of the human TOLL protein comprise a transmembrane domain.

10

In another embodiment, a TOLL of the present invention is identified based on the presence of a "leucine-rich repeat" domain in the protein or corresponding nucleic acid molecule. As used herein, the term "leucine-rich repeat" domain includes a protein domain having an amino acid sequence of about 50 amino acid residues and having a bit score for the alignment of the sequence to the leucine-rich repeat domain (LRR) of at least about 15. Preferably, a leucine-rich repeat domain includes at least about 20-60, more preferably about 20-50 amino acid residues, or about 20-40 amino acids and has a bit score for the alignment of the sequence to the leucine-rich repeat domain (LRR) of at least 20, 30, 40, 50 or greater. The leucine-rich repeat domain (LRR) has been assigned the PFAM Accession PF00560 (<http://genome.wustl.edu/Pfam/.html>). Leucine-rich repeat domains are described in, for example, Kobe, B. et al., (1994) *Trends in Biol. Sci.* 19:415-421, the contents of which are incorporated herein by reference.

15

20

25

30

35

40

45

50

15 To identify the presence of a leucine-rich repeat domain in a TOLL protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of HMMs (e.g., the Pfam database, release 2.1) using the default parameters (http://www.sanger.ac.uk/Software/Pfam/HMM_search). For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for MILPAT0063 and a score of 15 is the default threshold score for determining a hit. Alternatively, the threshold score for determining a hit can be lowered (e.g., to 8 bits). A description of the Pfam database can be found in Sonhammer et al. (1997) *Proteins* 28(3):405-420 and a detailed description of HMMs can be found, for example, in Gribskov et al. (1990) *Meth. Enzymol.* 183:146-159; Gribskov et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh et al. (1994) *J. Mol. Biol.* 235:1501-1531; and Stultz et al. (1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference. A search was performed against the HMM database resulting in the identification of a leucine-rich repeat domain in the amino acid sequence of human TOLL (SEQ ID NO:2) at about residues 58-105, 106-153, 154-201.

55

5

202-250, 251-297, and 298-345 of SEQ ID NO:2. The results of the search are set forth in Figure 2.

10

Accordingly, TOLL proteins having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a leucine-rich repeat domain of human TOLL (e.g., residues 58-105, 106-153, 154-201, 202-250, 251-297, or 298-345 of SEQ ID NO:2) or with a TM domain of human TOLL (e.g., residues 456-480 and 483-507 of SEQ ID NO:2) are within the scope of the invention.

15

In another embodiment, a TOLL molecule of the present invention is identified based on the presence of a β/α -class fold with α -helices on both faces of a parallel β sheet in the protein molecule. As used herein, the term " β/α -class fold with α -helices on both faces of a parallel β sheet" includes a domain having a β sheet structure in which the longer and more hydrophobic β strands are predicted to form interior staves in the beta sheet, while the more amphipathic β strands are at the edge of the sheet, with α -helices on both faces of the sheet (see Rock, F.L. et al. (1998) *Proc. Natl. Acad. Sci.*

20

25 U.S.A. 95: 588-593, the contents of which are incorporated herein by reference). An α helix is a rodlike coiled structure stabilized by hydrogen bonds between the NH and CO groups of the polypeptide chain, while a β sheet is a sheet comprised of different β strands which may be parallel or antiparallel, in which the strands are almost fully extended (as opposed to the tightly coiled α -helix); such structures are known in the art.

25

30 Isolated proteins of the present invention, preferably TOLL proteins, have an amino acid sequence sufficiently homologous to the amino acid sequence of SEQ ID NO:2 or are encoded by a nucleotide sequence sufficiently homologous to SEQ ID NO:1, 3, or 4. As used herein, the term "sufficiently homologous" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., 35 40 45 50 55 60 65 70 75 80 85 90 95 98 99 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995 3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995 4000 4005 4010 4015 4020 4025 4030 4035 4040 4045 4050 4055 4060 4065 4070 4075 4080 4085 4090 4095 4100 4105 4110 4115 4120 4125 4130 4135 4140 4145 4150 4155 4160 4165 4170 4175 4180 4185 4190 4195 4200 4205 4210 4215 4220 4225 4230 4235 4240 4245 4250 4255 4260 4265 4270 4275 4280 4285 4290 4295 4300 4305 4310 4315 4320 4325 4330 4335 4340 4345 4350 4355 4360 4365 4370 4375 4380 4385 4390 4395 4400 4405 4410 4415 4420 4425 4430 4435 4440 4445 4450 4455 4460 4465 4470 4475 4480 4485 4490 4495 4500 4505 4510 4515 4520 4525 4530 4535 4540 4545 4550 4555 4560 4565 4570 4575 4580 4585 4590 4595 4600 4605 4610 4615 4620 4625 4630 4635 4640 4645 4650 4655 4660 4665 4670 4675 4680 4685 4690 4695 4700 4705 4710 4715 4720 4725 4730 4735 4740 4745 4750 4755 4760 4765 4770 4775 4780 4785 4790 4795 4800 4805 4810 4815 4820 4825 4830 4835 4840 4845 4850 4855 4860 4865 4870 4875 4880 4885 4890 4895 4900 4905 4910 4915 4920 4925 4930 4935 4940 4945 4950 4955 4960 4965 4970 4975 4980 4985 4990 4995 5000 5005 5010 5015 5020 5025 5030 5035 5040 5045 5050 5055 5060 5065 5070 5075 5080 5085 5090 5095 5100 5105 5110 5115 5120 5125 5130 5135 5140 5145 5150 5155 5160 5165 5170 5175 5180 5185 5190 5195 5200 5205 5210 5215 5220 5225 5230 5235 5240 5245 5250 5255 5260 5265 5270 5275 5280 5285 5290 5295 5300 5305 5310 5315 5320 5325 5330 5335 5340 5345 5350 5355 5360 5365 5370 5375 5380 5385 5390 5395 5400 5405 5410 5415 5420 5425 5430 5435 5440 5445 5450 5455 5460 5465 5470 5475 5480 5485 5490 5495 5500 5505 5510 5515 5520 5525 5530 5535 5540 5545 5550 5555 5560 5565 5570 5575 5580 5585 5590 5595 5600 5605 5610 5615 5620 5625 5630 5635 5640 5645 5650 5655 5660 5665 5670 5675 5680 5685 5690 5695 5700 5705 5710 5715 5720 5725 5730 5735 5740 5745 5750 5755 5760 5765 5770 5775 5780 5785 5790 5795 5800 5805 5810 5815 5820 5825 5830 5835 5840 5845 5850 5855 5860 5865 5870 5875 5880 5885 5890 5895 5900 5905 5910 5915 5920 5925 5930 5935 5940 5945 5950 5955 5960 5965 5970 5975 5980 5985 5990 5995 6000 6005 6010 6015 6020 6025 6030 6035 6040 6045 6050 6055 6060 6065 6070 6075 6080 6085 6090 6095 6100 6105 6110 6115 6120 6125 6130 6135 6140 6145 6150 6155 6160 6165 6170 6175 6180 6185 6190 6195 6200 6205 6210 6215 6220 6225 6230 6235 6240 6245 6250 6255 6260 6265 6270 6275 6280 6285 6290 6295 6300 6305 6310 6315 6320 6325 6330 6335 6340 6345 6350 6355 6360 6365 6370 6375 6380 6385 6390 6395 6400 6405 6410 6415 6420 6425 6430 6435 6440 6445 6450 6455 6460 6465 6470 6475 6480 6485 6490 6495 6500 6505 6510 6515 6520 6525 6530 6535 6540 6545 6550 6555 6560 6565 6570 6575 6580 6585 6590 6595 6600 6605 6610 6615 6620 6625 6630 6635 6640 6645 6650 6655 6660 6665 6670 6675 6680 6685 6690 6695 6700 6705 6710 6715 6720 6725 6730 6735 6740 6745 6750 6755 6760 6765 6770 6775 6780 6785 6790 6795 6800 6805 6810 6815 6820 6825 6830 6835 6840 6845 6850 6855 6860 6865 6870 6875 6880 6885 6890 6895 6900 6905 6910 6915 6920 6925 6930 6935 6940 6945 6950 6955 6960 6965 6970 6975 6980 6985 6990 6995 7000 7005 7010 7015 7020 7025 7030 7035 7040 7045 7050 7055 7060 7065 7070 7075 7080 7085 7090 7095 7100 7105 7110 7115 7120 7125 7130 7135 7140 7145 7150 7155 7160 7165 7170 7175 7180 7185 7190 7195 7200 7205 7210 7215 7220 7225 7230 7235 7240 7245 7250 7255 7260 7265 7270 7275 7280 7285 7290 7295 7300 7305 7310 7315 7320 7325 7330 7335 7340 7345 7350 7355 7360 7365 7370 7375 7380 7385 7390 7395 7400 7405 7410 7415 7420 7425 7430 7435 7440 7445 7450 7455 7460 7465 7470 7475 7480 7485 7490 7495 7500 7505 7510 7515 7520 7525 7530 7535 7540 7545 7550 7555 7560 7565 7570 7575 7580 7585 7590 7595 7600 7605 7610 7615 7620 7625 7630 7635 7640 7645 7650 7655 7660 7665 7670 7675 7680 7685 7690 7695 7700 7705 7710 7715 7720 7725 7730 7735 7740 7745 7750 7755 7760 7765 7770 7775 7780 7785 7790 7795 7800 7805 7810 7815 7820 7825 7830 7835 7840 7845 7850 7855 7860 7865 7870 7875 7880 7885 7890 7895 7900 7905 7910 7915 7920 7925 7930 7935 7940 7945 7950 7955 7960 7965 7970 7975 7980 7985 7990 7995 8000 8005 8010 8015 8020 8025 8030 8035 8040 8045 8050 8055 8060 8065 8070 8075 8080 8085 8090 8095 8100 8105 8110 8115 8120 8125 8130 8135 8140 8145 8150 8155 8160 8165 8170 8175 8180 8185 8190 8195 8200 8205 8210 8215 8220 8225 8230 8235 8240 8245 8250 8255 8260 8265 8270 8275 8280 8285 8290 8295 8300 8305 8310 8315 8320 8325 8330 8335 8340 8345 8350 8355 8360 8365 8370 8375 8380 8385 8390 8395 8400 8405 8410 8415 8420 8425 8430 8435 8440 8445 8450 8455 8460 8465 8470 8475 8480 8485 8490 8495 8500 8505 8510 8515 8520 8525 8530 8535 8540 8545 8550 8555 8560 8565 8570 8575 8580 8585 8590 8595 8600 8605 8610 8615 8620 8625 8630 8635 8640 8645 8650 8655 8660 8665 8670 8675 8680 8685 8690 8695 8700 8705 8710 8715 8720 8725 8730 8735 8740 8745 8750 8755 8760 8765 8770 8775 8780 8785 8790 8795 8800 8805 8810 8815 8820 8825 8830 8835 8840 8845 8850 8855 8860 8865 8870 8875 8880 8885 8890 8895 8900 8905 8910 8915 8920 8925 8930 8935 8940 8945 8950 8955 8960 8965 8970 8975 8980 8985 8990 8995 9000 9005 9010 9015 9020 9025 9030 9035 9040 9045 9050 9055 9060 9065 9070 9075 9080 9085 9090 9095 9100 9105 9110 9115 9120 9125 9130 9135 9140 9145 9150 9155 9160 9165 9170 9175 9180 9185 9190 9195 9200 9205 9210 9215 9220 9225 9230 9235 9240 9245 9250 9255 9260 9265 9270 9275 9280 9285 9290 9295 9300 9305 9310 9315 9320 9325 9330 9335 9340 9345 9350 9355 9360 9365 9370 9375 9380 9385 9390 9395 9400 9405 9410 9415 9420 9425 9430 9435 9440 9445 9450 9455 9460 9465 9470 9475 9480 9485 9490 9495 9500 9505 9510 9515 9520 9525 9530 9535 9540 9545 9550 9555 9560 9565 9570 9575 9580 9585 9590 9595 9600 9605 9610 9615 9620 9625 9630 9635 9640 9645 9650 9655 9660 9665 9670 9675 9680 9685 9690 9695 9700 9705 9710 9715 9720 9725 9730 9735 9740 9745 9750 9755 9760 9765 9770 9775 9780 9785 9790 9795 9800 9805 9810 9815 9820 9825 9830 9835 9840 9845 9850 9855 9860 9865 9870 9875 9880 9885 9890 9895 9900 9905 9910 9915 9920 9925 9930 9935 9940 9945 9950 9955 9960 9965 9970 9975 9980 9985 9990 9995 10000 10005 10010 10015 10020 10025 10030 10035 10040 10045 10050 10055 10060 10065 10070 10075 10080 10085 10090 10095 10100 10105 10110 10115 10120 10125 10130 10135 10140 10145 10150 10155 10160 10165 10170 10175 10180 10185 10190 10195 10200 10205 10210 10215 10220 10225 10230 10235 10240 10245 10250 10255 10260 10265 10270 10275 10280 10285 10290 10295 10300 10305 10310 10315 10320 10325 10330 10335 10340 10345 10350 10355 10360 10365 10370 10375 10380 10385 10390 10395 10400 10405 10410 10415 10420 10425 10430 10435 10440

5

acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently homologous. Furthermore, amino acid or nucleotide sequences which share at least 30%, 40%, or 50%, preferably 60%, more preferably 70-80%, or 90-95% homology and share a common functional activity are defined herein as sufficiently homologous.

10

5 defined herein as sufficiently homologous.

15

As used interchangeably herein, an "TOLL activity", "biological activity of TOLL" or "functional activity of TOLL", refers to an activity exerted by a TOLL protein, polypeptide or nucleic acid molecule on a TOLL responsive cell or on a TOLL protein substrate, as determined *in vivo*, or *in vitro*, according to standard techniques. In 10 one embodiment, a TOLL activity is a direct activity, such as an association with a TOLL-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with which a TOLL protein binds or interacts in nature, such that TOLL-mediated function is achieved. A TOLL target molecule can be a non-TOLL molecule 20 or a TOLL protein or polypeptide of the present invention. In an exemplary 25 embodiment, a TOLL target molecule is a TOLL ligand. Alternatively, a TOLL activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the TOLL protein with a TOLL ligand.

30

Accordingly, another embodiment of the invention features isolated TOLL proteins and polypeptides having a TOLL activity. Preferred proteins are TOLL 35 proteins having at least one transmembrane domain, and, preferably, a TOLL activity. Other preferred proteins are TOLL proteins having at least one leucine-rich repeat domain and, preferably, a TOLL activity. Other preferred proteins are TOLL proteins 40 having at least one domain containing a β/α -class fold with α -helices on both faces of a parallel β sheet and, preferably, a TOLL activity. Additional preferred proteins have at 45 least one transmembrane domain and/or a leucine-rich repeat domain, and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, or 4.

45

The nucleotide sequence of the isolated human TOLL cDNA and the predicted 50 amino acid sequence of the human TOLL polypeptide are shown in Figure 1 and in SEQ ID NOs:1 and 2, respectively. The nucleotide sequence of the isolated partial mouse

55

- 13 -

5

TOLL DNA is shown in Figure 3. A plasmid containing the nucleotide sequence encoding human TOLL was deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on _____ and assigned Accession Number _____. A plasmid containing the nucleotide sequence

10

5 encoding partial mouse TOLL was deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on _____ and assigned Accession Number _____. These deposits will be maintained under the 15 terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. These deposits were made 20 merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

20

The human TOLL gene, which is approximately 2147 nucleotides in length, encodes a protein having a molecular weight of approximately 75 kD and which is approximately 548 amino acid residues in length. The partial mouse TOLL nucleotide 15 sequence is approximately 763 nucleotides in length.

30

Various aspects of the invention are described in further detail in the following subsections:

I. Isolated Nucleic Acid Molecules

35

20 One aspect of the invention pertains to isolated nucleic acid molecules that encode TOLL proteins or biologically active portions thereof, as well as nucleic acid 35 fragments sufficient for use as hybridization probes to identify TOLL-encoding nucleic acid molecules (e.g., TOLL mRNA) and fragments for use as PCR primers for the 40 amplification or mutation of TOLL nucleic acid molecules. As used herein, the term 45 25 "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

45

The term "isolated nucleic acid molecule" includes nucleic acid molecules which 30 are separated from other nucleic acid molecules which are present in the natural source 50 of the nucleic acid. For example, with regards to genomic DNA, the term "isolated"

5

includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated TOLL nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

10

15

20

25

30

35

40

45

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequence of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, as a hybridization probe, TOLL nucleic acid molecules can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd. ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can

5

be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to TOLL nucleotide sequences can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

10

In a preferred embodiment, an isolated nucleic acid molecule of the invention

15

5 comprises the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO: 4. The sequence of SEQ ID NO:1 corresponds to the human TOLL cDNA. This cDNA comprises sequences encoding the human TOLL protein (i.e., "the coding region", from nucleotides 92-1735), as well as 5' untranslated sequences (nucleotides 1-91) and 3' untranslated sequences (nucleotides 1736-2147). Alternatively, the nucleic acid

20

10 molecule can comprise only the coding region of SEQ ID NO:1 (e.g., nucleotides 92-1735, corresponding to SEQ ID NO:3). The sequence of SEQ ID NO:4 corresponds to a fragment of the mouse TOLL cDNA.

25

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide 15 sequence shown in SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, is one 30 which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that it can hybridize to the nucleotide 35 sequence shown in SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby forming a 40 stable duplex.

45

In still another preferred embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to the entire length of the nucleotide sequence shown in SEQ ID NO:1, 3, or 4, or the entire length of the 50 nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences.

55

5

Moreover, the nucleic acid molecule of the invention can comprise only a portion

of the nucleic acid sequence of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the

DNA insert of the plasmid deposited with ATCC as Accession Number _____, for

10

example, a fragment which can be used as a probe or primer or a fragment encoding a

5 portion of a TOLL protein, e.g., a biologically active portion of a TOLL protein. The

nucleotide sequence determined from the cloning of the TOLL gene allows for the

15

generation of probes and primers designed for use in identifying and/or cloning other

TOLL family members, as well as TOLL homologues from other species. The

probe/primer typically comprises substantially purified oligonucleotide. The

20

oligonucleotide typically comprises a region of nucleotide sequence that hybridizes

under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more

preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 consecutive nucleotides of a sense

25

sequence of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the

plasmid deposited with ATCC as Accession Number _____, of an anti-sense sequence

15 of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid

deposited with ATCC as Accession Number _____, or of a naturally occurring allelic

30

variant or mutant of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert

of the plasmid deposited with ATCC as Accession Number _____. In an exemplary

embodiment, a nucleic acid molecule of the present invention comprises a nucleotide

20 sequence which is greater than 491, 491-500, 500-550, 550-600, 600-650, 650-700, 700-

35

750, 750-800, 800-850, 850-900, 900-950, 950-1000, or more nucleotides in length and

hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ

40

ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited

with ATCC as Accession Number _____.

25 Probes based on the TOLL nucleotide sequences can be used to detect transcripts

or genomic sequences encoding the same or homologous proteins. In preferred

45

embodiments, the probe further comprises a label group attached thereto, e.g., the label

group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-

factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or

30 tissue which misexpress a TOLL protein, such as by measuring a level of a TOLL-

50

5

encoding nucleic acid in a sample of cells from a subject e.g., detecting TOLL mRNA levels or determining whether a genomic TOLL gene has been mutated or deleted.

10

A nucleic acid fragment encoding a "biologically active portion of a TOLL protein" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID

15

5 NO:1, 3, or 4 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. which encodes a polypeptide having a TOLL biological activity (the biological activities of the TOLL proteins are described herein). expressing the encoded portion of the TOLL protein (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the TOLL protein.

20

10 The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, due to degeneracy of the genetic code and thus encode the same TOLL proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1, 3, or 4, or the nucleotide

25

15 sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ 30 ID NO:2.

30

In addition to the TOLL nucleotide sequences shown in SEQ ID NO:1, 3, or 4, or 35 the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the TOLL proteins may exist within a population (e.g., the human population). Such genetic 40 polymorphism in the TOLL genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" 45 refer to nucleic acid molecules which include an open reading frame encoding a TOLL protein, preferably a mammalian TOLL protein, and can further include non-coding regulatory sequences, and introns.

45

Allelic variants of human TOLL include both functional and non-functional 50 TOLL proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the human TOLL protein that maintain the ability to bind a TOLL ligand

55

5

and/or to modulate signaling mechanisms involved in a TOLL related mechanism. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2 or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein.

10

5 Non-functional allelic variants are naturally occurring amino acid sequence variants of the human TOLL protein that do not have the ability to either bind a TOLL ligand and/or to modulate signaling mechanisms involved in, for example, an immune response. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion or premature truncation of the amino acid sequence 15 of SEQ ID NO:2 or a substitution, insertion or deletion in critical residues or critical 20 regions.

20

The present invention further provides non-human orthologues of the human TOLL protein. Orthologues of the human TOLL protein are proteins that are isolated 25 from non-human organisms and possess the same TOLL ligand binding properties as the 15 human TOLL protein. Orthologues of the human TOLL protein can readily be identified as comprising an amino acid sequence that is substantially homologous to 30 SEQ ID NO:2.

30

Moreover, nucleic acid molecules encoding other TOLL family members and, thus, which have a nucleotide sequence which differs from the TOLL sequences of SEQ 20 ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the 35 invention. For example, another TOLL cDNA can be identified based on the nucleotide sequence of human TOLL. Moreover, nucleic acid molecules encoding TOLL proteins from different species, and which, thus, have a nucleotide sequence which differs from 40 the TOLL sequences of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA 25 insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, a mouse TOLL cDNA can be identified based on the nucleotide sequence of a human TOLL.

40

Nucleic acid molecules corresponding to natural allelic variants and homologues 30 of the TOLL cDNAs of the invention can be isolated based on their homology to the 50 TOLL nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion

55

5

thercof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Nucleic acid molecules corresponding to natural allelic variants and homologues of the TOLL cDNAs of the invention can further be 10 isolated by mapping to the same chromosome or locus as the TOLL gene.

10

5 Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15, 20, 25, 30 or more nucleotides in length and hybridizes under 15 stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In other embodiment, the nucleic 20 acid is at least 30, 50, 100, 150, 200, 250, 253, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, or 950 nucleotides in length. As used herein, the term 25 "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other. Preferably, the conditions are 30 such that sequences at least about 70%, more preferably at least about 80%, even more 35 preferably at least about 85% or 90% homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization 40 conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50°C, preferably at 55°C, 45 more preferably at 60°C, and even more preferably at 65°C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1, 3, or 4 or corresponds to a naturally-occurring nucleic acid 50 molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

In addition to naturally-occurring allelic variants of the TOLL sequences that may exist in the population, the skilled artisan will further appreciate that changes can be 30 introduced by mutation into the nucleotide sequences of SEQ ID NO:1, 3, or 4, or the

50

55

- 20 -

5

nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. thereby leading to changes in the amino acid sequence of the encoded TOLL proteins, without altering the functional ability of the TOLL proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential"

10

5 amino acid residues can be made in the sequence of SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of TOLL (e.g., the sequence of SEQ ID NO:2) without altering the biological activity, whereas an "essential" amino acid residue is required for 10 biological activity. For example, amino acid residues that are conserved among the TOLL proteins of the present invention, e.g., those present in the transmembrane domain, are predicted to be particularly unamenable to alteration. Furthermore, additional amino acid residues that are conserved between the TOLL proteins of the present invention and 20 other members of the TOLL family are not likely to be amenable to alteration.

25

15 Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding TOLL proteins that contain changes in amino acid residues that are not essential for activity. Such TOLL proteins differ in amino acid sequence from SEQ ID NO:2, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein 30 comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 35 85%, 90%, 95%, 98% or more homologous to SEQ ID NO:2.

30

An isolated nucleic acid molecule encoding a TOLL protein homologous to the protein of SEQ ID NO:2 can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1, 3, or 40 25, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into SEQ 45 ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ by standard techniques, such as site-directed 30 mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A 50

55

5

"conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g.,

10

5 aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid

10 residue in a TOLL protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a TOLL coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for TOLL biological activity to

20 identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1, 3, or 4, or

25 15 the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

30

In a preferred embodiment, a mutant TOLL protein can be assayed for the ability to (1) interact with a non-TOLL protein molecule, e.g., a TOLL ligand; (2) activate a TOLL-dependent signal transduction pathway; or (3) modulate signaling mechanisms involved in, for example, an immune response.

35

In addition to the nucleic acid molecules encoding TOLL proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which

40 25 is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire TOLL coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is

45 30 antisense to a "coding region" of the coding strand of a nucleotide sequence encoding TOLL. The term "coding region" refers to the region of the nucleotide sequence

50

55

5

comprising codons which are translated into amino acid residues (e.g., the coding region of human TOLL corresponds to SEQ ID NO:3). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding TOLL. The term "noncoding region" refers to 5' and 3'

10

5 sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

15

Given the coding strand sequences encoding TOLL disclosed herein (e.g., SEQ ID NO:3), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be

20

10 complementary to the entire coding region of TOLL mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of TOLL mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of TOLL mRNA. An antisense

25

oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using

30

chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase

35

20 the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-

40

25 carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-

45

30 mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-

50

55

5 thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the
10 antisense nucleic acid can be produced biologically using an expression vector into which
15 5 a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from
the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of
interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered
20 to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA
10 and/or genomic DNA encoding a TOLL protein to thereby inhibit expression of the
protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by
25 conventional nucleotide complementarity to form a stable duplex, or, for example, in the
case of an antisense nucleic acid molecule which binds to DNA duplexes, through
specific interactions in the major groove of the double helix. An example of a route of
30 15 administration of antisense nucleic acid molecules of the invention include direct
injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified
to target selected cells and then administered systemically. For example, for systemic
35 administration, antisense molecules can be modified such that they specifically bind to
receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense
40 20 nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or
antigens. The antisense nucleic acid molecules can also be delivered to cells using the
vectors described herein. To achieve sufficient intracellular concentrations of the
45 25 antisense molecules, vector constructs in which the antisense nucleic acid molecule is
placed under the control of a strong pol II or pol III promoter are preferred.
30 30 In yet another embodiment, the antisense nucleic acid molecule of the invention
is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms
specific double-stranded hybrids with complementary RNA in which, contrary to the
45 35 usual β -units, the strands run parallel to each other (Gaultier et al. (1987) *Nucleic Acids
Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-
50 40 methylribonucleotide (Inoue et al. (1987) *Nucleic Acids Res.* 15:6131-6148) or a
chimeric RNA-DNA analogue (Inoue et al. (1987) *FEBS Lett.* 215:327-330).

5

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes

10

5 (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave TOLL mRNA transcripts to thereby inhibit translation of TOLL mRNA. A ribozyme having specificity for a TOLL-encoding nucleic acid can be

15

designed based upon the nucleotide sequence of a TOLL cDNA disclosed herein (i.e., SEQ ID NO:1, 3, or 4, or the nucleotide sequence of the DNA insert of the plasmid

20

10 deposited with ATCC as Accession Number ____). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a TOLL-encoding mRNA. See, e.g., Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S.

25

15 Patent No. 5,116,742. Alternatively, TOLL mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

30

Alternatively, TOLL gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the TOLL (e.g., the TOLL promoter and/or enhancers) to form triple helical structures that prevent transcription of 20 the TOLL gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. et al. (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

35

In yet another embodiment, the TOLL nucleic acid molecules of the present 40 invention can be modified at the base moiety, sugar moiety or phosphate backbone to 45 improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. et al. (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate 50 backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific

50

5

hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. et al. (1996) *supra*; Perry-O'Keefe et al. Proc. Natl. Acad. Sci. 93: 14670-675.

10

5 PNAS of TOLL nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAS can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAS of TOLL nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes. (e.g., S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. et al. (1996) *supra*; Perry-O'Keefe *supra*).

15

20

25

In another embodiment, PNAS of TOLL can be modified, (e.g., to enhance their

15 stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of TOLL nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (e.g., RNase H and

20 DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of

25 PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. et al. (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine

30 phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. et al. (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a

35 stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. et al. (1996) *supra*). Alternatively, chimeric molecules can be

50

55

5

synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. et al. (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

10

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating

15

5 transport across the cell membrane (see, e.g., Letsinger et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, e.g., Krol et al. (1988) *Bio-Techniques* 6:958-976) or

10 intercalating agents. (See, e.g., Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

20

25

II. Isolated TOLL Proteins and Anti-TOLL Antibodies

30

15 One aspect of the invention pertains to isolated TOLL proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-TOLL antibodies. In one embodiment, native TOLL proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, TOLL proteins are

20 produced by recombinant DNA techniques. Alternative to recombinant expression, a TOLL protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

35

40 An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or

45 25 tissue source from which the TOLL protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of TOLL protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of

30 30 cellular material" includes preparations of TOLL protein having less than about 30% (by dry weight) of non-TOLL protein (also referred to herein as a "contaminating protein"),

50

55

5 more preferably less than about 20% of non-TOLL protein, still more preferably less than about 10% of non-TOLL protein, and most preferably less than about 5% non-TOLL protein. When the TOLL protein or biologically active portion thereto is
10 recombinantly produced, it is also preferably substantially free of culture medium, i.e.,
5 culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

15 The language "substantially free of chemical precursors or other chemicals" includes preparations of TOLL protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one
20 embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of TOLL protein having less than about 30% (by dry weight) of chemical precursors or non-TOLL chemicals, more preferably less than about 20% chemical precursors or non-TOLL chemicals, still more preferably less than about 10% chemical precursors or non-TOLL chemicals, and most preferably less than about 5%
25 chemical precursors or non-TOLL chemicals.

30 As used herein, a "biologically active portion" of a TOLL protein includes a fragment of a TOLL protein which participates in an interaction between a TOLL molecule and a non-TOLL molecule. Biologically active portions of a TOLL protein include peptides comprising amino acid sequences sufficiently homologous to or derived
35 from the amino acid sequence of the TOLL protein, e.g., the amino acid sequence shown in SEQ ID NO:2, which include less amino acids than the full length TOLL proteins, and exhibit at least one activity of a TOLL protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the TOLL protein. A biologically active portion of a TOLL protein can be a polypeptide which is, for
40 example, 10, 25, 50, 100, 200 or more amino acids in length. Biologically active portions of a TOLL protein can be used as targets for developing agents which modulate a TOLL mediated activity.

45 In one embodiment, a biologically active portion of a TOLL protein comprises at least one transmembrane domain, and/or at least one leucine-rich repeat domain and/or
30 at least one domain containing a β/α -class fold with α -helices on both faces of a parallel β sheet. It is to be understood that a preferred biologically active portion of a TOLL

5

protein of the present invention may contain at least one transmembrane domain. Another preferred biologically active portion of a TOLL protein may contain at least one leucine-rich repeat domain. Another preferred biologically active portion of a TOLL 10 protein may contain at least one domain containing a β/α -class fold with α -helices on 5 both faces of a parallel β sheet. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and 15 evaluated for one or more of the functional activities of a native TOLL protein.

10

In a preferred embodiment, the TOLL protein has an amino acid sequence shown 20 in SEQ ID NO:2. In other embodiments, the TOLL protein is substantially homologous to SEQ ID NO:2, and retains the functional activity of the protein of SEQ ID NO:2, yet 25 differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the TOLL protein is a protein which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to SEQ 15 ID NO:2.

30

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps 35 can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for 40 comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 730%, preferably at least 40%, more 45 preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (e.g., when aligning a second sequence to the TOLL amino acid sequence of SEQ ID NO:2 having 548 amino 50 acid residues, at least 250, preferably at least 300, more preferably at least 350, even more preferably at least 400, and even more preferably at least 450 or 500 amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid

55

5

"homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of

10

the two sequences.

15

5 The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at 10 <http://www.gcg.com>), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at 15 <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 20 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the 25 algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue 30 table, a gap length penalty of 12 and a gap penalty of 4.

35

20 The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the 25 NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to TOLL nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain 30 amino acid sequences homologous to TOLL protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) *Nucleic Acids Res.* 25(17):3389-3402. When

40

45

50

55

5

utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

10

The invention also provides TOLL chimeric or fusion proteins. As used herein, a TOLL "chimeric protein" or "fusion protein" comprises a TOLL polypeptide

15

5 operatively linked to a non-TOLL polypeptide. A "TOLL polypeptide" refers to a polypeptide having an amino acid sequence corresponding to TOLL, whereas a "non-TOLL polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the TOLL protein, e.g., a protein which is different from the TOLL protein and which is derived from the

20

10 same or a different organism. Within a TOLL fusion protein the TOLL polypeptide can correspond to all or a portion of a TOLL protein. In a preferred embodiment, a TOLL fusion protein comprises at least one biologically active portion of a TOLL protein. In another preferred embodiment, a TOLL fusion protein comprises at least two biologically active portions of a TOLL protein. Within the fusion protein, the term

25

15 "operatively linked" is intended to indicate that the TOLL polypeptide and the non-TOLL polypeptide are fused in-frame to each other. The non-TOLL polypeptide can be fused to the N-terminus or C-terminus of the TOLL polypeptide.

30

For example, in one embodiment, the fusion protein is a GST-TOLL fusion protein in which the TOLL sequences are fused to the C-terminus of the GST sequences.

20 Such fusion proteins can facilitate the purification of recombinant TOLL.

35

In another embodiment, the fusion protein is a TOLL protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of TOLL can be increased through use of a heterologous signal sequence.

40

25 The TOLL fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The TOLL fusion proteins can be used to affect the bioavailability of a TOLL substrate. Use of TOLL fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a TOLL protein; 30 (ii) mis-regulation of the TOLL gene; and (iii) aberrant post-translational modification of a TOLL protein.

50

5

Moreover, the TOLL-fusion proteins of the invention can be used as immunogens to produce anti-TOLL antibodies in a subject, to purify TOLL ligands and in screening assays to identify molecules which inhibit the interaction of TOLL with a TOLL substrate.

10

5 Preferably, a TOLL chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with 15 conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini,

20

10 filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene 25 can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene

25

15 fragments which can subsequently be annealed and reamplified to generate a chimeric 30 gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A TOLL- 35 encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the TOLL protein.

35

The present invention also pertains to variants of the TOLL proteins which 40 function as either TOLL agonists (mimetics) or as TOLL antagonists. Variants of the TOLL proteins can be generated by mutagenesis, e.g., discrete point mutation or truncation of a TOLL protein. An agonist of the TOLL proteins can retain substantially 45 the same, or a subset, of the biological activities of the naturally occurring form of a TOLL protein. An antagonist of a TOLL protein can inhibit one or more of the activities of the naturally occurring form of the TOLL protein by, for example, competitively modulating a TOLL-mediated activity of a TOLL protein. Thus, specific 50 biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological

50

55

5

activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the TOLL protein.

10

In one embodiment, variants of a TOLL protein which function as either TOLL agonists (mimetics) or as TOLL antagonists can be identified by screening

15

5 combinatorial libraries of mutants, e.g., truncation mutants, of a TOLL protein for TOLL protein agonist or antagonist activity. In one embodiment, a variegated library of TOLL variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of TOLL variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides

20

10 into gene sequences such that a degenerate set of potential TOLL sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of TOLL sequences therein. There are a variety of methods which can be used to produce libraries of potential TOLL variants

25

25 from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential TOLL sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura et al. (1984)

30

20 *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477.

35

In addition, libraries of fragments of a TOLL protein coding sequence can be used to generate a variegated population of TOLL fragments for screening and subsequent selection of variants of a TOLL protein. In one embodiment, a library of 40 coding sequence fragments can be generated by treating a double stranded PCR fragment of a TOLL coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from 45 different nicked products, removing single stranded portions from reformed duplexes by 30 treatment with S1 nuclease, and ligating the resulting fragment library into an expression

50

55

5

vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the TOLL protein.

10

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA

15

libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of TOLL proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting

20

library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in

25

combination with the screening assays to identify TOLL variants (Arkin and Yourvan 1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave et al. (1993) *Protein Engineering* 6(3):327-331).

30

In one embodiment, cell based assays can be exploited to analyze a variegated TOLL library. For example, a library of expression vectors can be transfected into a cell line which ordinarily responds to a particular ligand in a TOLL-dependent manner. The

35

transfected cells are then contacted with the ligand and the effect of expression of the mutant on signaling by the ligand can be detected, e.g., by measuring the activity of a TOLL-regulated transcription factor. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the ligand, and the individual clones further characterized.

40

25 An isolated TOLL protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind TOLL using standard techniques for polyclonal and monoclonal antibody preparation. A full-length TOLL protein can be used or, alternatively, the invention provides antigenic peptide fragments of TOLL for use as immunogens. The antigenic peptide of TOLL comprises at least 8 amino acid

45

30 residues of the amino acid sequence shown in SEQ ID NO:2 and encompasses an epitope of TOLL such that an antibody raised against the peptide forms a specific

50

55

5

immune complex with TOLL. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

10

5 Preferred epitopes encompassed by the antigenic peptide are regions of TOLL that are located on the surface of the protein, e.g., hydrophilic regions, as well as regions with high antigenicity.

15

A TOLL immunogen typically is used to prepare antibodies by immunizing a suitable subject, (e.g., rabbit, goat, mouse or other mammal) with the immunogen. An 10 appropriate immunogenic preparation can contain, for example, recombinantly expressed TOLL protein or a chemically synthesized TOLL polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with 20 an immunogenic TOLL preparation induces a polyclonal anti-TOLL antibody response.

25

15 Accordingly, another aspect of the invention pertains to anti-TOLL antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and 30 immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as TOLL. Examples of immunologically active portions of immunoglobulin 20 molecules include F(ab) and F(ab')₂ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and 35 monoclonal antibodies that bind TOLL. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of 40 immunoreacting with a particular epitope of TOLL. A monoclonal antibody composition thus typically displays a single binding affinity for a particular TOLL protein with which it immunoreacts.

45

Polyclonal anti-TOLL antibodies can be prepared as described above by 30 immunizing a suitable subject with a TOLL immunogen. The anti-TOLL antibody titer in the immunized subject can be monitored over time by standard techniques, such as 50 with an enzyme linked immunosorbent assay (ELISA) using immobilized TOLL. If

55

5 desired, the antibody molecules directed against TOLL can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, e.g., when the anti-TOLL antibody titers are highest, antibody-producing cells can be

10 5 obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497 (see also, Brown et al. (1981) *J. Immunol.* 127:539-46; Brown et al. (1980) *J. Biol. Chem.* 255:4980-83; Yeh et al. (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh et al. (1982) *Int. J. Cancer* 29:269-75), the

15 10 more recent human B cell hybridoma technique (Kozbor et al. (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole et al. (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally R. H. 20 25 Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 15 15 54:387-402; M. L. Gefter et al. (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) 30 30 from a mammal immunized with a TOLL immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a 20 hybridoma producing a monoclonal antibody that binds TOLL.

35 Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-TOLL 40 40 monoclonal antibody (see, e.g., G. Galfré et al. (1977) *Nature* 266:55052; Gefter et al. *Somatic Cell Genet.*, cited *supra*; Lerner, *Yale J. Biol. Med.*, cited *supra*; Kenneth, 45 45 25 *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (e.g., a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made 50 30 30 by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing

hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, e.g., the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind TOLL, e.g., using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-TOLL antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (e.g., an antibody phage display library) with TOLL to thereby isolate immunoglobulin library members that bind TOLL.

Kits for generating and screening phage display libraries are commercially available (e.g., the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner et al. U.S. Patent No. 5,223,409; Kang et al. PCT International Publication No. WO 92/18619; Dower et al. PCT International Publication No. WO 91/17271; Winter et al. PCT International Publication WO 92/20791; Markland et al. PCT International Publication No. WO 92/15679; Breitling et al. PCT International Publication WO 93/01288; McCafferty et al. PCT International Publication No. WO 92/01047; Garrard et al. PCT International Publication No. WO 92/09690; Ladner et al. PCT International Publication No. WO 90/02809; Fuchs et al. (1991) *Bio/Technology* 9:1370-1372; Hay et al. (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse et al. (1989) *Science* 246:1275-1281; Griffiths et al. (1993) *EMBO J* 12:725-734; Hawkins et al. (1992) *J. Mol. Biol.* 226:889-896; Clarkson et al. (1991) *Nature* 352:624-628; Gram et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrad et al. (1991) *Bio/Technology* 9:1373-1377;

5

Hoogenboom et al. (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty et al. *Nature* (1990) 348:552-554.

10

Additionally, recombinant anti-TOLL antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions,

15

5 which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson et al. International Application No. PCT/US86/02269; Akira, et al. European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496;

20

10 Morrison et al. European Patent Application 173,494; Neuberger et al. PCT International Publication No. WO 86/01533; Cabilly et al. U.S. Patent No. 4,816,567; Cabilly et al. European Patent Application 125,023; Better et al. (1988) *Science* 240:1041-1043; Liu et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu et al. (1987) *J. Immunol.*

25

15 139:3521-3526; Sun et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura et al. (1987) *Canc. Res.* 47:999-1005; Wood et al. (1985) *Nature* 314:446-449; and Shaw et al. (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison, S. L. (1985) *Science* 229:1202-1207; Oi et al. (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones et al. (1986) *Nature* 321:552-525; Verhoeven et al. (1988) *Science* 239:1534; and Beidler et al. (1988) *J. Immunol.* 141:4053-4060.

30

20 An anti-TOLL antibody (e.g., monoclonal antibody) can be used to isolate TOLL proteins by standard techniques, such as affinity chromatography or

35

immunoprecipitation. An anti-TOLL antibody can facilitate the purification of natural TOLL proteins from cells and of recombinantly produced TOLL proteins expressed in host cells. Moreover, an anti-TOLL antibody can be used to detect TOLL protein (e.g.,

40

25 in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the TOLL protein. Anti-TOLL antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance.

45

30 Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive

50

55

5

10

15

20

25

30

35

40

45

50

55

materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

10 III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a TOLL protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

5

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel: *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., TOLL proteins, mutant forms of TOLL proteins, fusion proteins, and the like).

10

15

20

25

30

35

40

45

50

55

The recombinant expression vectors of the invention can be designed for expression of TOLL proteins in prokaryotic or eukaryotic cells. For example, TOLL proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

5

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion

10

5 vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from

15

10 the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase.

20

Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRITS (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST),

25

15 maltose E binding protein, or protein A, respectively, to the target recombinant protein.

30

Purified fusion proteins can be utilized in TOLL activity assays, (e.g., direct assays or competitive assays described in detail below), or to generate antibodies specific for TOLL proteins, for example. In a preferred embodiment, a TOLL fusion protein expressed in a retroviral expression vector of the present invention can be 20 utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (e.g., six (6) weeks).

35

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann et al., (1988) *Gene* 69:301-315) and pET 11d (Studier et al., *Gene*

40

25 *Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gnl0-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gnl). This viral 45 polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident

50

55

5

prophage harboring a T7 *gnl* gene under the transcriptional control of the lacUV 5 promoter.

10

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada et al., (1992) *Nucleic Acids Res.* 20:2111-2118).

15

Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

20

In another embodiment, the TOLL expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYEpSec1 (Baldari, et al., (1987) *EMBO J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al., (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

25

Alternatively, TOLL proteins can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith et al. (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

30

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman et al. (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements.

35

For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd. ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989.

40

45

50

55

5 In another embodiment, the recombinant mammalian expression vector is
capable of directing expression of the nucleic acid preferentially in a particular cell type
(e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-
10 specific regulatory elements are known in the art. Non-limiting examples of suitable
5 tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al.
(1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988)
15 *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winioto and
Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji et al. (1983) *Cell*
33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters
20 (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA*
86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) *Science* 230:912-916),
and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No.
25 4,873,316 and European Application Publication No. 264,166). Developmentally-
regulated promoters are also encompassed, for example the murine hox promoters
15 (Kessel and Gruss (1990) *Science* 249:374-379) and the α -fetoprotein promoter
(Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

30 The invention further provides a recombinant expression vector comprising a
DNA molecule of the invention cloned into the expression vector in an antisense-
orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in
20 a manner which allows for expression (by transcription of the DNA molecule) of an
RNA molecule which is antisense to TOLL mRNA. Regulatory sequences operatively
linked to a nucleic acid cloned in the antisense orientation can be chosen which direct
the continuous expression of the antisense RNA molecule in a variety of cell types, for
35 instance viral promoters and/or enhancers, or regulatory sequences can be chosen which
allow for direct constitutive, tissue specific or cell type specific expression of antisense RNA. The
40 antisense expression vector can be in the form of a recombinant plasmid, phagemid or
attenuated virus in which antisense nucleic acids are produced under the control of a
45 high efficiency regulatory region, the activity of which can be determined by the cell
type into which the vector is introduced. For a discussion of the regulation of gene
30 expression using antisense genes see Weintraub, H. et al., Antisense RNA as a molecular
tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

5

Another aspect of the invention pertains to host cells into which a TOLL nucleic acid molecule of the invention is introduced, e.g., a TOLL nucleic acid molecule within a recombinant expression vector or a TOLL nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome.

10

5 The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in 15 succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the 20 scope of the term as used herein.

20

A host cell can be any prokaryotic or eukaryotic cell. For example, a TOLL protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or 25 mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

25

15 Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized 30 techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated 20 transfection, lipofection, or electroporation. Suitable methods for transforming or 35 transfecting host cells can be found in Sambrook, et al. (*Molecular Cloning: A Laboratory Manual*, 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

40

For stable transfection of mammalian cells, it is known that, depending upon the 25 expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these 45 integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred 30 selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be 50 introduced into a host cell on the same vector as that encoding a TOLL protein or can be

55

5 introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

10 A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) a TOLL protein. Accordingly, the invention further provides methods for producing a TOLL protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a TOLL protein has been introduced) in a suitable medium such that a TOLL protein is produced. In another 15 embodiment, the method further comprises isolating a TOLL protein from the medium or the host cell.

20 The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which TOLL-coding sequences have been 25 introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous TOLL sequences have been introduced into their genome or homologous recombinant animals in which endogenous TOLL sequences have been 30 altered. Such animals are useful for studying the function and/or activity of a TOLL and for identifying and/or evaluating modulators of TOLL activity. As used herein, a 35 "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, 40 dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops 45 and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous TOLL gene has been 50 altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

5

A transgenic animal of the invention can be created by introducing a TOLL-encoding nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The TOLL cDNA sequence of SEQ ID NO:1 can

10

5 be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human TOLL gene, such as a mouse or rat TOLL gene, or the sequence of SEQ ID NO:4 can be used as a transgene. Alternatively, a TOLL gene homologue, such as another TOLL family member, can be isolated based on hybridization to the TOLL cDNA sequences of SEQ ID NO:1, 3, or 4, or the DNA insert

10 of the plasmid deposited with ATCC as Accession Number _____ (described further in subsection I above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a

20

25 TOLL transgene to direct expression of a TOLL protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder et al., U.S. Patent No. 4,873,191 by Wagner et al. and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar 30 methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a TOLL transgene in its genome and/or expression of TOLL mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a TOLL protein can further 40 be bred to other transgenic animals carrying other transgenes.

35

45 To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a TOLL gene into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the TOLL gene. The TOLL gene can be a human gene (e.g., the cDNA of SEQ ID NO:1), but more preferably, is a non- 50 human homologue of a human TOLL gene (e.g., a cDNA isolated by stringent hybridization with the nucleotide sequence of SEQ ID NO:1, or the sequence of SEQ ID

40

55

5 NO: 4). For example, a mouse TOLL gene can be used to construct a homologous
recombination nucleic acid molecule, e.g., a vector, suitable for altering an endogenous
TOLL gene in the mouse genome. In a preferred embodiment, the homologous
10 recombination nucleic acid molecule is designed such that, upon homologous
recombination, the endogenous TOLL gene is functionally disrupted (i.e., no longer
encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the
15 homologous recombination nucleic acid molecule can be designed such that, upon
homologous recombination, the endogenous TOLL gene is mutated or otherwise altered
but still encodes functional protein (e.g., the upstream regulatory region can be altered to
20 thereby alter the expression of the endogenous TOLL protein). In the homologous
recombination nucleic acid molecule, the altered portion of the TOLL gene is flanked at
its 5' and 3' ends by additional nucleic acid sequence of the TOLL gene to allow for
homologous recombination to occur between the exogenous TOLL gene carried by the
25 homologous recombination nucleic acid molecule and an endogenous TOLL gene in a
cell, e.g., an embryonic stem cell. The additional flanking TOLL nucleic acid sequence
is of sufficient length for successful homologous recombination with the endogenous
30 gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are
included in the homologous recombination nucleic acid molecule (see, e.g., Thomas,
K.R. and Capecchi, M. R. (1987) *Cell* 51:503 for a description of homologous
35 recombination vectors). The homologous recombination nucleic acid molecule is
introduced into a cell, e.g., an embryonic stem cell line (e.g., by electroporation) and
cells in which the introduced TOLL gene has homologously recombined with the
endogenous TOLL gene are selected (see e.g., Li, E. et al. (1992) *Cell* 69:915). The
40 selected cells can then be injected into a blastocyst of an animal (e.g., a mouse) to form
aggregation chimeras (see e.g., Bradley, A. in *Teratocarcinomas and Embryonic Stem
45 Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A
chimeric embryo can then be implanted into a suitable pseudopregnant female foster
animal and the embryo brought to term. Progeny harboring the homologously
recombined DNA in their germ cells can be used to breed animals in which all cells of
50 the animal contain the homologously recombined DNA by germline transmission of the
transgene. Methods for constructing homologous recombination nucleic acid molecules,

5

e.g., vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec et al.; WO 91/01140 by Smithies et al.; WO 92/0968 by Zijlstra et al.; and WO 93/04169 by Berns et al.

10

5 In another embodiment, transgenic non-humans animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the 10 FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al. (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of 15 "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

15

20

25

30

35

40

45

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. et al. (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, 20 a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species 25 from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyst and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, e.g., the somatic cell, is isolated.

IV. Pharmaceutical Compositions

50

The TOLL nucleic acid molecules, fragments of TOLL proteins, and anti-TOLL antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such

55

5 compositions typically comprise the nucleic acid molecule, protein, or antibody and a
pharmaceutically acceptable carrier. As used herein the language "pharmaceutically
acceptable carrier" is intended to include any and all solvents, dispersion media,
10 coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents,
15 and the like, compatible with pharmaceutical administration. The use of such media and
agents for pharmaceutically active substances is well known in the art. Except insofar as
any conventional media or agent is incompatible with the active compound, use thereof
in the compositions is contemplated. Supplementary active compounds can also be
incorporated into the compositions.

20 10 A pharmaceutical composition of the invention is formulated to be compatible
with its intended route of administration. Examples of routes of administration include
parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation),
transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions
25 used for parenteral, intradermal, or subcutaneous application can include the following
15 components: a sterile diluent such as water for injection, saline solution, fixed oils,
polyethylene glycols, glycerine, propylene glycol or other synthetic solvents;
antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as
30 ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic
acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of
20 tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases,
such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be
35 enclosed in ampoules, disposable syringes or multiple dose vials made of glass or
plastic.

40 40 Pharmaceutical compositions suitable for injectable use include sterile aqueous
solutions (where water soluble) or dispersions and sterile powders for the
extemporaneous preparation of sterile injectable solutions or dispersion. For
45 intravenous administration, suitable carriers include physiological saline, bacteriostatic
water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In
all cases, the composition must be sterile and should be fluid to the extent that easy
30 50 syringability exists. It must be stable under the conditions of manufacture and storage
and must be preserved against the contaminating action of microorganisms such as

5 bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the

10 5 maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various 15 antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include 20 isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be 25 brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

30 Sterile injectable solutions can be prepared by incorporating the active compound (e.g., a fragment of a TOLL protein or an anti-TOLL antibody) in the 35 required amount in an appropriate solvent with one or a combination of ingredients 40 enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains 45 a basic dispersion medium and the required other ingredients from those enumerated 50 above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a 55 previously sterile-filtered solution thereof.

60 Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral 65 therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. oral compositions can also be prepared 70 using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically 75 compatible binding agents, and/or adjuvant materials can be included as part of the 80 composition. The tablets, pills, capsules, troches and the like can contain any of the 85 following ingredients, or compounds of a similar nature: a binder such as

90

5 microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint,
10 5 methyl salicylate, or orange flavoring.

15 For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

20 Systemic administration can also be by transmucosal or transdermal means. For 10 transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active
25 15 compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

30 The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

35 20 In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, 40 25 polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova
45 45 Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled
50 30 in the art, for example, as described in U.S. Patent No. 4,522,811.

5

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

10

10 Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio

15

15 LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

20

20 The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the

25

40 method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high

45

30 performance liquid chromatography.

30

35

40

45

50

55

5

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see e.g., Chen et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-

10

5 The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g., retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene

15

10 delivery system.

20

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

25

V. Uses and Methods of the Invention

15 The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (e.g., therapeutic and prophylactic). As described herein, a TOLL protein of the invention has one or more of

30

20 the following activities: (1) it interacts with a non-TOLL protein molecule, e.g., a TOLL ligand; (2) it activates a TOLL-dependent signal transduction pathway; and (3) it modulates signaling mechanisms involved in, for example, an immune response, and, thus, can be used to, for example, (1) modulate the interaction with a non-TOLL protein molecule; (2) to activate a TOLL-dependent signal transduction pathway; and (3) to

40

25 modulate signalling mechanisms involved in, for example, an immune response.

35

The isolated nucleic acid molecules of the invention can be used, for example, to express TOLL protein (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect TOLL mRNA (e.g., in a biological sample) or a genetic alteration in a TOLL gene, and to modulate TOLL activity, as described further below.

45

30 The TOLL proteins can be used to treat disorders characterized by insufficient or excessive production of a TOLL substrate or production of TOLL inhibitors. In

50

55

5 addition, the TOLL proteins can be used to screen for naturally occurring TOLL substrates, to screen for drugs or compounds which modulate TOLL activity, as well as to treat disorders characterized by insufficient or excessive production of TOLL protein or production of TOLL protein forms which have decreased, aberrant or unwanted

10 5 activity compared to TOLL wild type protein. Moreover, the anti-TOLL antibodies of the invention can be used to detect and isolate TOLL proteins, regulate the bioavailability of TOLL proteins, and modulate TOLL activity.

15

20 A. Screening Assays:

10 The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, i.e., candidate or test compounds or agents (e.g., peptides, peptidomimetics, small molecules or other drugs) which bind to TOLL proteins, have a stimulatory or inhibitory effect on, for example, TOLL expression or TOLL activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of

25 15 TOLL substrate.

30 In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a TOLL protein or polypeptide or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a TOLL protein

35 20 or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library

40 25 methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

45 Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994) *J. Med.*

5 *Chem.* 37:2678; Cho et al. (1993) *Science* 261:1303; Carrell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop et al. (1994) *J. Med. Chem.* 37:1233.

10 Libraries of compounds may be presented in solution (e.g., Houghten (1992)

5 *Biotechniques* 13:412-421); or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull et al. (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla et al. (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra*).

20 In one embodiment, an assay is a cell-based assay in which a cell which expresses a TOLL protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate TOLL activity is determined. Determining the ability of the test compound to modulate TOLL activity

25 15 can be accomplished by monitoring, for example, the activity of a TOLL-regulated transcription factor. The cell, for example, can be of mammalian origin.

30 The ability of the test compound to modulate TOLL binding to a substrate or to bind to TOLL can also be determined. Determining the ability of the test compound to modulate TOLL binding to a substrate can be accomplished, for example, by coupling 20 the TOLL substrate with a radioisotope or enzymatic label such that binding of the TOLL substrate to TOLL can be determined by detecting the labeled TOLL substrate in a complex. Determining the ability of the test compound to bind TOLL can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to TOLL can be determined by detecting the 40 25 labeled TOLL compound in a complex. For example, compounds (e.g., TOLL substrates) can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label 30 35 detected by determination of conversion of an appropriate substrate to product.

5 It is also within the scope of this invention to determine the ability of a compound (e.g., a TOLL substrate) to interact with TOLL without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with TOLL without the labeling of either the compound or the TOLL.

10 5 McConnell, H. M. et al. (1992) *Science* 257:1906-1912. As used herein, a 15 "microphysiometer" (e.g., Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and TOLL.

20 10 In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a TOLL target molecule (e.g., a TOLL substrate) with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the TOLL target molecule. Determining the ability of the test compound to modulate the activity of a TOLL target molecule can be accomplished, for example, 25 15 by determining the ability of the TOLL protein to bind to or interact with the TOLL target molecule.

30 Determining the ability of the TOLL protein or a biologically active fragment thereof, to bind to or interact with a TOLL target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred 35 20 embodiment, determining the ability of the TOLL protein to bind to or interact with a TOLL target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by 40 detecting induction of a cellular second messenger of the target (e.g., intracellular Ca^{2+} , diacylglycerol, IP_3 , and the like), detecting catalytic/enzymatic activity of the target an 45 25 appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, e.g., luciferase), or detecting a target-regulated cellular response.

50 45 In yet another embodiment, an assay of the present invention is a cell-free assay in which a TOLL protein or biologically active portion thereof is contacted with a test 50 30 compound and the ability of the test compound to bind to the TOLL protein or biologically active portion thereof is determined. Preferred biologically active portions

5 of the TOLL proteins to be used in assays of the present invention include fragments
which participate in interactions with non-TOLL molecules, e.g., fragments with high
surface probability scores. Binding of the test compound to the TOLL protein can be
10 determined either directly or indirectly as described above. In a preferred embodiment,
5 the assay includes contacting the TOLL protein or biologically active portion thereof
with a known compound which binds TOLL to form an assay mixture, contacting the
15 assay mixture with a test compound, and determining the ability of the test compound to
interact with a TOLL protein, wherein determining the ability of the test compound to
interact with a TOLL protein comprises determining the ability of the test compound to
20 preferentially bind to TOLL or biologically active portion thereof as compared to the
known compound.

25 In another embodiment, the assay is a cell-free assay in which a TOLL protein or
biologically active portion thereof is contacted with a test compound and the ability of
the test compound to modulate (e.g., stimulate or inhibit) the activity of the TOLL
30 protein or biologically active portion thereof is determined. Determining the ability of
the test compound to modulate the activity of a TOLL protein can be accomplished, for
example, by determining the ability of the TOLL protein to bind to a TOLL target
molecule by one of the methods described above for determining direct binding.
Determining the ability of the TOLL protein to bind to a TOLL target molecule can also
35 be accomplished using a technology such as real-time Biomolecular Interaction Analysis
(BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo
et al. (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology
40 for studying biospecific interactions in real time, without labeling any of the interactants
(e.g., BIACore). Changes in the optical phenomenon of surface plasmon resonance
45 (SPR) can be used as an indication of real-time reactions between biological molecules.

45 In an alternative embodiment, determining the ability of the test compound to
modulate the activity of a TOLL protein can be accomplished by determining the ability
of the TOLL protein to further modulate the activity of a downstream effector of a
TOLL target molecule. For example, the activity of the effector molecule on an
50 appropriate target can be determined or the binding of the effector to an appropriate
target can be determined as previously described.

5 In yet another embodiment, the cell-free assay involves contacting a TOLL protein or biologically active portion thereof with a known compound which binds the TOLL protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the TOLL.

10 5 protein, wherein determining the ability of the test compound to interact with the TOLL protein comprises determining the ability of the TOLL protein to preferentially bind to or modulate the activity of a TOLL target molecule.

15 10 The cell-free assays of the present invention are amenable to use of both soluble and/or membrane-bound forms of isolated proteins (e.g., TOLL proteins or biologically active portions thereof). In the case of cell-free assays in which a membrane-bound form of an isolated protein is used it may be desirable to utilize a solubilizing agent such that the membrane-bound form of the isolated protein is maintained in solution.

20 25 Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)_n, 3-[(3-cholamidopropyl)dimethylammonio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

20 30 20 In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either TOLL or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a TOLL protein, or interaction of a TOLL protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/ TOLL fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are

5 then combined with the test compound or the test compound and either the non-adsorbed target protein or TOLL protein, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following 10 incubation, the beads or microtitre plate wells are washed to remove any unbound 15 components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of TOLL binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the 20 screening assays of the invention. For example, either a TOLL protein or a TOLL target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated TOLL protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (e.g., biotinylation kit, Pierce 25 Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well 30 plates (Pierce Chemical). Alternatively, antibodies reactive with TOLL protein or target molecules but which do not interfere with binding of the TOLL protein to its target molecule can be derivatized to the wells of the plate, and unbound target or TOLL protein trapped in the wells by antibody conjugation. Methods for detecting such 35 complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the TOLL protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the TOLL protein or target molecule.

In another embodiment, modulators of TOLL expression are identified in a 40 method wherein a cell is contacted with a candidate compound and the expression of 45 TOLL mRNA or protein in the cell is determined. The level of expression of TOLL mRNA or protein in the presence of the candidate compound is compared to the level of expression of TOLL mRNA or protein in the absence of the candidate compound. The 50 candidate compound can then be identified as a modulator of TOLL expression based on this comparison. For example, when expression of TOLL mRNA or protein is greater 30 (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of TOLL mRNA or

5 protein expression. Alternatively, when expression of TOLL mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of TOLL mRNA or protein expression. The level of TOLL mRNA or protein expression in the cells can be
10 5 determined by methods described herein for detecting TOLL mRNA or protein.

In yet another aspect of the invention, the TOLL proteins can be used as "bait
15 proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No.
5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.*
268:12046-12054; Bartel et al. (1993) *Biotechniques* 14:920-924; Iwabuchi et al. (1993)
10 *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind
20 to or interact with TOLL ("TOLL-binding proteins" or "TOLL-bp") and are involved in
TOLL activity. Such TOLL-binding proteins are also likely to be involved in the
propagation of signals by the TOLL proteins or TOLL targets as, for example,
25 downstream elements of a TOLL-mediated signaling pathway. Alternatively, such
15 TOLL-binding proteins are likely to be TOLL inhibitors.

The two-hybrid system is based on the modular nature of most transcription
30 factors, which consist of separable DNA-binding and activation domains. Briefly, the
assay utilizes two different DNA constructs. In one construct, the gene that codes for a
TOLL protein is fused to a gene encoding the DNA binding domain of a known
35 transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a
library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is
fused to a gene that codes for the activation domain of the known transcription factor. If
40 the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a TOLL-
dependent complex, the DNA-binding and activation domains of the transcription factor
45 are brought into close proximity. This proximity allows transcription of a reporter gene
(e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the
transcription factor. Expression of the reporter gene can be detected and cell colonies
containing the functional transcription factor can be isolated and used to obtain the
cloned gene which encodes the protein which interacts with the TOLL protein.

5 In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a TOLL 10 protein can be confirmed *in vivo*, e.g., in an animal such as an animal model for a TOLL 15 associated disorder.

This invention further pertains to novel agents identified by the above-described 15 screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a TOLL modulating agent, an antisense TOLL 20 nucleic acid molecule, a TOLL-specific antibody, or a TOLL-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, 25 this invention pertains to uses of novel agents identified by the above-described 15 screening assays for treatments as described herein.

30 **B. Detection Assays**

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as 20 polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with 35 genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

40 **25 1. Chromosome Mapping**

Once the sequence (or a portion of the sequence) of a gene has been isolated, this 45 sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the TOLL 30 nucleotide sequences, described herein, can be used to map the location of the TOLL

5

genes on a chromosome. The mapping of the TOLL sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

10

Briefly, TOLL genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the TOLL nucleotide sequences. Computer

15

5 analysis of the TOLL sequences can be used to predict primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene

corresponding to the TOLL sequences will yield an amplified fragment.

20

10 Somatic cell hybrids are prepared by fusing somatic cells from different mammals (e.g., human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but human cells can, the one human chromosome that contains

25

15 the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. (D'Eustachio P. et al. (1983) *Science* 220:919-924). Somatic cell hybrids

30

20 containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

35

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the TOLL nucleotide sequences to 40

25 design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a TOLL sequence to its chromosome include *in situ* hybridization 45 (described in Fan, Y. et al. (1990) *Proc. Natl. Acad. Sci. USA*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to 30 chromosome specific cDNA libraries.

50

55

5

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been

10

blocked in metaphase by a chemical such as colcemid that disrupts the mitotic spindle.

15

5 The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity

10 for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma et al., *Human Chromosomes: A Manual of Basic Techniques* (Pergamon Press, New York 1988).

20

25

Reagents for chromosome mapping can be used individually to mark a single

15 chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

30

35

20 Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, *Mendelian Inheritance in Man*, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can

25 then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland, J. et al. (1987) *Nature*, 325:783-787.

40

45

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the TOLL gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected

30 individuals, then the mutation is likely to be the causative agent of the particular disease.

Comparison of affected and unaffected individuals generally involves first looking for

50

55

5

structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

10

5

2. Tissue Typing

15

The TOLL sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

20

25

30

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the TOLL nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

35

40

45

50

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The TOLL nucleotide sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared

55

5 for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 can comfortably provide positive individual
10 identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding
5 amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ
ID NO:3 are used, a more appropriate number of primers for positive individual
15 identification would be 500-2,000.

If a panel of reagents from TOLL nucleotide sequences described herein is used
20 to generate a unique identification database for an individual, those same reagents can
10 later be used to identify tissue from that individual. Using the unique identification
database, positive identification of the individual, living or dead, can be made from
25 extremely small tissue samples.

3. Use of Partial TOLL Sequences in Forensic Biology

25 15 DNA-based identification techniques can also be used in forensic biology.
Forensic biology is a scientific field employing genetic typing of biological evidence
30 found at a crime scene as a means for positively identifying, for example, a perpetrator
of a crime. To make such an identification, PCR technology can be used to amplify
DNA sequences taken from very small biological samples such as tissues, e.g., hair or
20 skin, or body fluids, e.g., blood, saliva, or semen found at a crime scene. The amplified
35 sequence can then be compared to a standard, thereby allowing identification of the
origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide
40 reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can
25 enhance the reliability of DNA-based forensic identifications by, for example, providing
another "identification marker" (i.e. another DNA sequence that is unique to a particular
45 individual). As mentioned above, actual base sequence information can be used for
identification as an accurate alternative to patterns formed by restriction enzyme
generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 are
30 particularly appropriate for this use as greater numbers of polymorphisms occur in the
50 noncoding regions, making it easier to differentiate individuals using this technique.

5

Examples of polynucleotide reagents include the TOLL nucleotide sequences or portions thereof, e.g., fragments derived from the noncoding regions of SEQ ID NO:1 having a length of at least 20 bases, preferably at least 30 bases.

10

The TOLL nucleotide sequences described herein can further be used to provide

15

5 polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, e.g., brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such TOLL probes can be used to identify tissue by species and/or by organ type.

20

10 In a similar fashion, these reagents, e.g., TOLL primers or probes can be used to screen tissue culture for contamination (i.e. screen for the presence of a mixture of different types of cells in a culture).

25

15 C. Predictive Medicine:

30

15 The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining TOLL protein and/or nucleic acid expression as well as TOLL activity, in

35

20 the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant or unwanted TOLL expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with TOLL protein, nucleic acid

40

25 expression or activity. For example, mutations in a TOLL gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with TOLL protein, nucleic acid expression or activity.

45

Another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of TOLL in clinical trials.

50

These and other agents are described in further detail in the following sections.

55

5

1. Diagnostic Assays

10

An exemplary method for detecting the presence or absence of TOLL protein or nucleic acid in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting TOLL protein or nucleic acid (e.g., mRNA, genomic DNA) that encodes TOLL protein such that the presence of TOLL protein or nucleic acid is detected in the biological sample. A preferred agent for detecting TOLL mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to TOLL mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length TOLL nucleic acid, such as the nucleic acid of SEQ ID NO:1, 3, or 4, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to TOLL mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

15

15 A preferred agent for detecting TOLL protein is an antibody capable of binding to TOLL protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect TOLL mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of TOLL mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of TOLL protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and

20

25

30

35

40

45

50

55

5 immunofluorescence. *In vitro* techniques for detection of TOLL genomic DNA include
10 Southern hybridizations. Furthermore, *in vivo* techniques for detection of TOLL protein
15 include introducing into a subject a labeled anti-TOLL antibody. For example, the
antibody can be labeled with a radioactive marker whose presence and location in a
5 subject can be detected by standard imaging techniques.

10 In one embodiment, the biological sample contains protein molecules from the
15 test subject. Alternatively, the biological sample can contain mRNA molecules from the
test subject or genomic DNA molecules from the test subject. A preferred biological
sample is a serum sample isolated by conventional means from a subject.

20 10 In another embodiment, the methods further involve obtaining a control
biological sample from a control subject, contacting the control sample with a
25 compound or agent capable of detecting TOLL protein, mRNA, or genomic DNA, such
that the presence of TOLL protein, mRNA or genomic DNA is detected in the biological
sample, and comparing the presence of TOLL protein, mRNA or genomic DNA in the
15 control sample with the presence of TOLL protein, mRNA or genomic DNA in the test
sample.

30 The invention also encompasses kits for detecting the presence of TOLL in a
biological sample. For example, the kit can comprise a labeled compound or agent
35 capable of detecting TOLL protein or mRNA in a biological sample; means for
20 determining the amount of TOLL in the sample; and means for comparing the amount of
TOLL in the sample with a standard. The compound or agent can be packaged in a
suitable container. The kit can further comprise instructions for using the kit to detect
40 TOLL protein or nucleic acid.

45 2. Prognostic Assays

25 25 The diagnostic methods described herein can furthermore be utilized to identify
subjects having or at risk of developing a disease or disorder associated with aberrant or
unwanted TOLL expression or activity. As used herein, the term "aberrant" includes a
45 TOLL expression or activity which deviates from the wild type TOLL expression or
activity. Aberrant expression or activity includes increased or decreased expression or
30 activity, as well as expression or activity which does not follow the wild type
developmental pattern of expression or the subcellular pattern of expression. For

5 example, aberrant TOLL expression or activity is intended to include the cases in which
a mutation in the TOLL gene causes the TOLL gene to be under-expressed or over-
expressed and situations in which such mutations result in a non-functional TOLL
10 protein or a protein which does not function in a wild-type fashion, e.g., a protein which
5 does not interact with a TOLL ligand or one which interacts with a non-TOLL ligand.
As used herein, the term "unwanted" includes an unwanted phenomenon involving a
15 TOLL associated response, such as an immune response. For example, the term
unwanted includes a TOLL expression or activity which is undesirable in a subject.

The assays described herein, such as the preceding diagnostic assays or the
20 10 following assays, can be utilized to identify a subject having or at risk of developing a
disorder associated with a misregulation in TOLL protein activity or nucleic acid
expression, such as, for example, an immune disorder. Alternatively, the prognostic
assays can be utilized to identify a subject having or at risk for developing a disorder
25 associated with a misregulation in TOLL protein activity or nucleic acid expression.
25 Thus, the present invention provides a method for identifying a disease or disorder
associated with aberrant or unwanted TOLL expression or activity in which a test
sample is obtained from a subject and TOLL protein or nucleic acid (e.g., mRNA or
30 genomic DNA) is detected, wherein the presence of TOLL protein or nucleic acid is
diagnostic for a subject having or at risk of developing a disease or disorder associated
20 with aberrant or unwanted TOLL expression or activity. As used herein, a "test sample"
35 refers to a biological sample obtained from a subject of interest. For example, a test
sample can be a biological fluid (e.g., serum), cell sample, or tissue.

40 Furthermore, the prognostic assays described herein can be used to determine
whether a subject can be administered an agent (e.g., an agonist, antagonist,
45 25 peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate)
to treat a disease or disorder associated with aberrant or unwanted TOLL expression or
activity. For example, such methods can be used to determine whether a subject can be
effectively treated with an agent for a TOLL associated disorder. Thus, the present
invention provides methods for determining whether a subject can be effectively treated
30 with an agent for a disorder associated with aberrant or unwanted TOLL expression or
activity in which a test sample is obtained and TOLL protein or nucleic acid expression
50

5

or activity is detected (e.g., wherein the abundance of TOLL protein or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant or unwanted TOLL expression or activity).

10

The methods of the invention can also be used to detect genetic alterations in a

15

5 TOLL gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in TOLL protein activity or nucleic acid expression. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a TOLL-protein, or

10 the mis-expression of the TOLL gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a TOLL gene; 2) an addition of one or more nucleotides to a TOLL gene; 3) a substitution of one or more nucleotides of a TOLL gene, 4) a chromosomal

20

25 rearrangement of a TOLL gene; 5) an alteration in the level of a messenger RNA transcript of a TOLL gene, 6) aberrant modification of a TOLL gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a TOLL gene, 8) a non-wild type level of a TOLL-protein, 9) allelic loss of a TOLL gene, and 10) inappropriate post-translational modification of a TOLL-protein. As described herein, there are a large number of assays

30 known in the art which can be used for detecting alterations in a TOLL gene. A preferred biological sample is a tissue or serum sample isolated by conventional means from a subject.

35

35 In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a 40 ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) *Science* 241:1077-1080; and Nakazawa et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the TOLL gene (see Abravaya et al. (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of

45 50 30 collecting a sample of cells from a subject, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or

50

55

5 more primers which specifically hybridize to a TOLL gene under conditions such that hybridization and amplification of the TOLL gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the
10 amplification product and comparing the length to a control sample. It is anticipated
5 that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

15 Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. et al., (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. et al., (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-
20 1177), Q-Beta Replicase (Lizardi, P.M. et al. (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such
25 molecules are present in very low numbers.

15 In an alternative embodiment, mutations in a TOLL gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel
30 electrophoresis and compared. Differences in fragment length sizes between sample and
20 control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score
35 for the presence of specific mutations by development or loss of a ribozyme cleavage site.

40 In other embodiments, genetic mutations in TOLL can be identified by
25 hybridizing a sample and control nucleic acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin, M.T. et al. (1996) *Human Mutation* 7: 244-255; Kozal, M.J. et al. (1996) *Nature Medicine* 2: 753-
45 759). For example, genetic mutations in TOLL can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, M.T. et al. *supra*.
30 Briefly, a first hybridization array of probes can be used to scan through long stretches
50 of DNA in a sample and control to identify base changes between the sequences by

5

making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is
10 5 composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

15

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the TOLL gene and detect mutations by comparing the sequence of the sample TOLL with the corresponding wild-type (control) 10 sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the 20 25 diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, e.g., PCT International Publication No. WO 94/16101; Cohen et al. (1996) *Adv. Chromatogr.* 36:127-162; and Griffin et al. (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

30

Other methods for detecting mutations in the TOLL gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA 20 35 or RNA/DNA heteroduplexes (Myers et al. (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type TOLL sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex 40 45 such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched 50 regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation.

50

5

See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

10

In still another embodiment, the mismatch cleavage reaction employs one or 5 more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in TOLL cDNAs obtained from samples of cells. For example, the *mutY* enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a TOLL 10 sequence, e.g., a wild-type TOLL sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

15

15 In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in TOLL genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility 20 between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA* 86:2766, see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992)

25

25 Genet. Anal. Tech. Appl. 9:73-79). Single-stranded DNA fragments of sample and control TOLL nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting 30 alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of 35 the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the 40 subject method utilizes heteroduplex analysis to separate double stranded heteroduplex 45 molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

50

55

5

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

10

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the

15

known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl Acad. Sci USA* 86:6230). Such allele specific

20

oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

25

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention.

30

Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

35

40

45

50

55

5 The methods described herein may be performed, for example, by utilizing pre-
10 packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent
 described herein, which may be conveniently used, e.g., in clinical settings to diagnose
15 patients exhibiting symptoms or family history of a disease or illness involving a TOLL
5 gene.

15 Furthermore, any cell type or tissue in which TOLL is expressed may be utilized
 in the prognostic assays described herein.

20 3. Monitoring of Effects During Clinical Trials

20 10 Monitoring the influence of agents (e.g., drugs) on the expression or activity of a
 TOLL protein (e.g., the modulation of signaling mechanisms involved in immunity) can
 be applied not only in basic drug screening, but also in clinical trials. For example, the
25 effectiveness of an agent determined by a screening assay as described herein to increase
 TOLL gene expression, protein levels, or upregulate TOLL activity, can be monitored in
 clinical trials of subjects exhibiting decreased TOLL gene expression, protein levels, or
30 downregulated TOLL activity. Alternatively, the effectiveness of an agent determined
 by a screening assay to decrease TOLL gene expression, protein levels, or downregulate
 TOLL activity, can be monitored in clinical trials of subjects exhibiting increased TOLL
 gene expression, protein levels, or upregulated TOLL activity. In such clinical trials, the
35 20 expression or activity of a TOLL gene, and preferably, other genes that have been
 implicated in, for example, a TOLL-associated disorder can be used as a "read out" or
 markers of the phenotype of a particular cell.

40 For example, and not by way of limitation, genes, including TOLL, that are
 modulated in cells by treatment with an agent (e.g., compound, drug or small molecule)
45 25 which modulates TOLL activity (e.g., identified in a screening assay as described
 herein) can be identified. Thus, to study the effect of agents on TOLL-associated
 disorders (e.g., disorders characterized by deregulated signaling mechanisms involved in
 immunity), for example, in a clinical trial, cells can be isolated and RNA prepared and
 analyzed for the levels of expression of TOLL and other genes implicated in the TOLL-
 associated disorder, respectively. The levels of gene expression (e.g., a gene expression
50 30 pattern) can be quantified by northern blot analysis or RT-PCR, as described herein, or

5

alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of TOLL or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined 10 before, and at various points during treatment of the individual with the agent.

10

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i)

20

10 obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a TOLL protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the TOLL protein,

25

mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of 15 expression or activity of the TOLL protein, mRNA, or genomic DNA in the pre-administration sample with the TOLL protein, mRNA, or genomic DNA in the post-administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be 30 desirable to increase the expression or activity of TOLL to higher levels than detected,

35

20 i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of TOLL to lower levels than detected, i.e. to decrease the effectiveness of the agent. According to such an embodiment, TOLL expression or activity may be used as an indicator of the 35 effectiveness of an agent, even in the absence of an observable phenotypic response.

40

25

D. Methods of Treatment:

45

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant or unwanted TOLL expression or activity. With regards to both 30 prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics.

50

55

5 "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of
10 how a patient's genes determine his or her response to a drug (e.g., a patient's "drug response phenotype", or "drug response genotype".) Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the TOLL molecules of the present invention or TOLL modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who
15 10 will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

20 1. Prophylactic Methods

25 In one aspect, the invention provides a method for preventing in a subject, a
30 15 disease or condition associated with an aberrant or unwanted TOLL expression or activity, by administering to the subject a TOLL protein or an agent which modulates TOLL expression or at least one TOLL activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted TOLL expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as
35 20 described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the TOLL aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of TOLL aberrancy, for example, a TOLL protein, TOLL agonist or TOLL antagonist
40 25 agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

45 2. Therapeutic Methods

30 Another aspect of the invention pertains to methods of modulating TOLL expression or activity for therapeutic purposes. Accordingly, in an exemplary 35 embodiment, the modulatory method of the invention involves contacting a cell with a TOLL protein or agent that modulates one or more of the activities of TOLL protein
50

5 activity associated with the cell. An agent that modulates TOLL protein activity can be
10 an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring
target molecule of a TOLL protein (e.g., a TOLL substrate), a TOLL antibody, a TOLL
agonist or antagonist, a peptidomimetic of a TOLL agonist or antagonist, or other small
15 molecule. In one embodiment, the agent stimulates one or more TOLL activities.
Examples of such stimulatory agents include active TOLL protein and a nucleic acid
20 molecule encoding TOLL that has been introduced into the cell. In another
embodiment, the agent inhibits one or more TOLL activities. Examples of such
inhibitory agents include antisense TOLL nucleic acid molecules, anti-TOLL antibodies,
25 and TOLL inhibitors. These modulatory methods can be performed *in vitro* (e.g., by
culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the
agent to a subject). As such, the present invention provides methods of treating an
individual afflicted with a disease or disorder characterized by aberrant or unwanted
30 expression or activity of a TOLL protein or nucleic acid molecule (e.g., rheumatoid
15 arthritis, systemic lupus erythematosus, myasthenia gravis, Grave's disease, Sjogren
syndrome, polymyositis and dermatomyositis, psoriasis, pemphigus vulgaris, bullous
35 pemphigoid, inflammatory bowel disease, Kawasaki disease, asthma, and graft vs. host
disease). In one embodiment, the method involves administering an agent (e.g., an agent
identified by a screening assay described herein), or combination of agents that
40 modulates (e.g., upregulates or downregulates) TOLL expression or activity. In another
embodiment, the method involves administering a TOLL protein or nucleic acid
molecule as therapy to compensate for reduced, aberrant, or unwanted TOLL expression
45 or activity.

40 Stimulation of TOLL activity is desirable in situations in which TOLL is
45 abnormally downregulated and/or in which increased TOLL activity is likely to have a
50 beneficial effect. For example, stimulation of TOLL activity is desirable in situations in
which a TOLL is downregulated and/or in which increased TOLL activity is likely to
have a beneficial effect. Likewise, inhibition of TOLL activity is desirable in situations
in which TOLL is abnormally upregulated and/or in which decreased TOLL activity is
55 likely to have a beneficial effect.

5

3. Pharmacogenomics

10

The TOLL molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on TOLL activity (e.g., TOLL gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) TOLL-associated disorders associated with aberrant or unwanted TOLL activity. In conjunction with such treatment, pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a TOLL molecule or TOLL modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a TOLL molecule or TOLL modulator.

15

15 Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. et al. (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11):983-985 and Linder, M.W. et al. (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

20

20 One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (e.g., a "biallelic" gene marker map which consists of 60,000-100,000 polymorphic or variable

25

25 30 35 40 45 50

55

5

sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a

10

5 high resolution map can be generated from a combination of some ten-million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-

15

10 associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

20

15 Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drug's target is known (e.g., a TOLL protein of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a 30 particular drug response.

30

35 As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of 40 genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation 45 as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive 50 metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently

50

55

5

experience exaggerated drug response and side effects when they receive standard doses.

10

If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who

5 do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

15

Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (e.g., a TOLL molecule or TOLL modulator of the present 10 invention) can give an indication whether gene pathways related to toxicity have been turned on.

20

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to 15 dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a TOLL molecule or TOLL modulator, such as a modulator identified by one of the exemplary screening assays described herein.

30

This invention is further illustrated by the following examples which should not 20 be construed as limiting. The contents of all references, patents and published patent 35 applications cited throughout this application are incorporated herein by reference.

EXAMPLES

40

25 EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF HUMAN AND MOUSE TOLL cDNA

45

In this example, the identification and characterization of the gene encoding human TOLL (clone jthKa089g09) and mouse TOLL (clone jtmBa212b08) is described.

30

50

55

5

Isolation of the Human TOLL cDNA

10

The invention is based, at least in part, on the discovery of a human gene encoding a novel protein, referred to herein as TOLL. A clone was originally identified based on sequence homology to Toll proteins, and based on the sequence of this first 5 clone, primers were designed and used to screen a human fat cell library (obtained from Clonetech). A positive human clone, jthKa089g09, was identified. The entire sequence of the human clone was determined and found to contain an open reading frame termed 15 human "TOLL".

15

20

The nucleotide sequence encoding the human TOLL protein is shown in Figure 1 10 and is set forth as SEQ ID NO:1. The full length protein encoded by this nucleic acid comprises about 548 amino acids and has the amino acid sequence shown in Figure 1 and set forth as SEQ ID NO:2. The coding region (open reading frame) of SEQ ID NO:1 is set forth as SEQ ID NO:3. Clone jthKa089g09, comprising the entire coding 25 region of the human TOLL gene, was deposited with the American Type Culture 15 Collection (ATCC®), 10801 University Boulevard, Manassas, VA 20110-2209, on _____, and assigned Accession No. _____.

25

30

Isolation of the Mouse TOLL cDNA

40

35

The invention is further based, at least in part, on the discovery of a partial 20 mouse nucleic acid sequence, referred to herein as mouse TOLL nucleic acid molecule. A clone was originally identified based on sequence homology to Toll proteins, and 35 based on the sequence of this first clone, primers were designed and used to screen a mouse cell library. A positive mouse clone, jtmBa212b08, was identified. The sequence of this mouse clone was determined and is termed mouse "TOLL".

45

25 A partial nucleotide sequence of the mouse TOLL gene is shown in Figure 3 and is set forth as SEQ ID NO:4. Clone jtmBa212b08, comprising the entire sequence of 40 SEQ ID NO:4, was deposited with the American Type Culture Collection (ATCC®), 10801 University Boulevard, Manassas, VA 20110-2209, on _____, and assigned Accession No. _____.

30

50

55

5

Analysis of the Human TOLL Molecule

A BLASTX 1.4.9 search, using a score of 100 and a word length of 3 (Altschul et al. (1990) *J. Mol. Biol.* 215:403) of the translated nucleotide sequence of human

10

TOLL revealed that human TOLL is similar to the rat MEGF5 protein (Accession

15

5 Number AB011531), the neurogenic extracellular slit protein slit2 (Accession Number AF055585), human slit-2 protein (Accession Number AB017168), human slit-3 protein (Accession Number AB017169), human slit-1 protein (Accession Number AB017167) and the rat MEGF4 protein (Accession Number AB014462). The human TOLL protein is 39% identical to the rat MEGF5 protein (Accession Number AB011531) over

20

10 translated nucleotides 191 to 643, 30% identical to this sequence over translated nucleotides 188 to 715, 28% identical to this sequence over translated nucleotides 347 to 787, 30% identical to this sequence over translated nucleotides 842 to 1189, 39% identical to this sequence over translated nucleotides 914 to 1189, 27% identical to this sequence over translated nucleotides 842 to 1276, 35% identical over translated

25

15 nucleotides 320 to 643, 37% identical over translated nucleotides 482 to 715, and 36% identical over translated nucleotides 839 to 946. The human TOLL protein is 37% identical to the neurogenic extracellular slit protein slit2 (Accession Number AF055585) over translated nucleotides 191 to 643, 30% identical to this sequence over translated nucleotides 188 to 730, 38% identical to this sequence over translated nucleotides 911 to

30

20 1189, 28% identical to this sequence over translated nucleotides 842 to 1246, 26% identical to this sequence over translated nucleotides 347 to 775, 36% identical to this sequence over translated nucleotides 311 to 586, 30% identical to this sequence over translated nucleotides 338 to 715, 28% identical to this sequence over residues 842 to 1258, 26% identical to this sequence over translated nucleotides 410 to 805, 22%

35

40 25 identical to this sequence over translated nucleotides 842 to 1276, 31% identical to this sequence over translated nucleotides 320 to 643, 25% identical to this sequence over translated nucleotides 275 to 601, 35% identical to this sequence over translated nucleotides 272 to 514, 29% identical to this sequence over translated nucleotides 419 to 790, 27% identical to this sequence over translated nucleotides 386 to 715, 30%

45

30 identical to this sequence over translated nucleotides 914 to 1168, 28% identical to this sequence over translated nucleotides 839 to 1156, 28% identical to this sequence over

50

5

translated nucleotides 275 to 571, 28% identical to this sequence over translated nucleotides 851 to 1168, 27% identical to this sequence over translated nucleotides 635 to 1090, 29% identical to this sequence over translated nucleotides 491 to 793, 26% identical to this sequence over translated nucleotides 263 to 601, is 27% identical to this sequence over translated nucleotides 458 to 787, 26% identical to this sequence over translated nucleotides 842 to 1162, 23% identical to this sequence over translated nucleotides 695 to 1084, 35% identical to this sequence over translated nucleotides 479 to 709, 30% identical to this sequence over translated nucleotides 551 to 802, 25% identical to this sequence over translated nucleotides 896 to 1162, 38% identical to this sequence over translated nucleotides 275 to 436, 25% identical to this sequence over translated nucleotides 383 to 643, 28% identical to this sequence over translated nucleotides 839 to 1090, and 30% identical to this sequence over translated nucleotides 281 to 508.

10

15

20

25

30

35

40

45

50

A BLASTN 1.4.9 search, using a score of 100 and a word length of 12 (Altschul et al. (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human TOLL revealed that the TOLL sequence is similar to the mouse mRNA for Knowles Solter mouse 2 cell *Mus musculus* cDNA clone 1124974 (Accession Number AA692768), to Soares infant brain 1NIB *Homo sapiens* cDNA clone mRNA (Accession Number R54798), to Soares infant brain 1NIB *Homo sapiens* cDNA clone mRNA (Accession Number H12046) and to Soares infant brain 1NIB *Homo sapiens* cDNA clone mRNA (Accession Number H12045). The TOLL nucleic acid molecule is 92% identical to the mouse mRNA for Knowles Solter mouse 2 cell *Mus musculus* cDNA clone 1124974 (Accession Number AA692768) over nucleotides 1038 to 548. The TOLL nucleic acid molecule is 100% identical to Soares infant brain 1NIB *Homo sapiens* cDNA clone mRNA (Accession Number R54798) over nucleotides 2061 to 1709, and 1711 to 1664, is 84% identical to this same sequence over nucleotides 1641-1603, and is 88% identical to this sequence over nucleotides 1656-1622. The TOLL nucleic acid molecule is 99% identical to Soares infant brain 1NIB *Homo sapiens* cDNA clone mRNA (Accession Number H12046) over nucleotides 2060 to 1850, is 98% identical to this sequence over nucleotides 1859-1709, is 75% identical to this sequence over nucleotides 1711-1658, is 92% identical to this sequence over nucleotides 1629-1603 and is 62% identical to this

5 sequence over nucleotides 1685-1622. The TOLL nucleic acid molecule is 98% identical to Soares infant brain INIB Homo sapiens cDNA clone mRNA (Accession Number H12045) over nucleotides 1323-1535, is 73% identical to this sequence over 10 nucleotides 1584 to 1680, is 67% identical to this sequence over nucleotides 1516-1614, 15 is 100% identical to this sequence over nucleotides 1576-1615, is 94% identical to this sequence over nucleotides 1686 to 1719, is 82% identical to this sequence over nucleotides 1663 to 1697, and is 66% identical to this sequence over nucleotides 1582 to 1629.

20 10 Tissue Distribution of TOLL mRNA

25 This Example describes the tissue distribution of TOLL mRNA, as can be determined by Northern blot hybridization and *in situ* hybridization.

30 Northern blot hybridizations with the various RNA samples are performed under standard conditions and washed under stringent conditions, i.e., 0.2xSSC at 65°C. The 15 DNA probe is radioactively labeled with ^{32}P -dCTP using the Prime-It kit (Stratagene, La Jolla, CA) according to the instructions of the supplier. Filters containing human or mouse mRNA (MultiTissue Northern I and MultiTissue Northern II from Clontech, Palo Alto, CA) are probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

35 20 For *in situ* analysis, various tissues obtained from brains, e.g. rat or monkey brains, are first frozen on dry ice. Ten-micrometer-thick coronal sections of the tissues are postfixed with 4% formaldehyde in DEPC treated 1X phosphate- buffered saline at room temperature for 10 minutes before being rinsed twice in DEPC 1X phosphate- buffered saline and once in 0.1 M triethanolamine-HCl (pH 8.0). Following incubation 40 25 in 0.25% acetic anhydride-0.1 M triethanolamine-HCl for 10 minutes, sections are rinsed in DEPC 2X SSC (1X SSC is 0.15M NaCl plus 0.015M sodium citrate). Tissue is then dehydrated through a series of ethanol washes, incubated in 100% chloroform for 5 minutes, and then rinsed in 100% ethanol for 1 minute and 95% ethanol for 1 minute and allowed to air dry.

5 Hybridizations are performed with ^{35}S -radiolabeled ($5 \times 10^7 \text{ cpm/ml}$) cRNA probes. Probes are incubated in the presence of a solution containing 600 mM NaCl, 10 mM Tris (pH 7.5), 1 mM EDTA, 0.01% sheared salmon sperm DNA, 0.01% yeast 10 tRNA, 0.05% yeast total RNA type X1, 1 X Denhardt's solution, 50% formamide, 10% 15 dextran sulfate, 100 mM dithiothreitol, 0.1% sodium dodecyl sulfate (SDS), and 0.1% sodium thiosulfate for 18 hours at 55°C.

15 After hybridization, slides are washed with 2 X SSC. Sections are then sequentially incubated at 37°C in TNE (a solution containing 10 mM Tris-HCl (pH 7.6), 20 500 mM NaCl, and 1 mM EDTA), for 10 minutes, in TNE with 10 μg of RNase A per ml for 30 minutes, and finally in TNE for 10 minutes. Slides are then rinsed with 2 X SSC 25 at room temperature, washed with 2 X SSC at 50°C for 1 hour, washed with 0.2 X SSC at 55°C for 1 hour, and 0.2 X SSC at 60°C for 1 hour. Sections are then dehydrated rapidly through serial ethanol-0.3 M sodium acetate concentrations before being air dried and exposed to Kodak Biomax MR scientific imaging film for 24 hours and 30 15 subsequently dipped in NB-2 photoemulsion and exposed at 4°C for 7 days before being developed and counter stained.

35

**EXAMPLE 2: EXPRESSION OF RECOMBINANT TOLL PROTEIN IN
 BACTERIAL CELLS**

20 In this example, TOLL protein is expressed as a recombinant glutathione-S- 35 transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, the TOLL sequence is fused to GST and this fusion 40 polypeptide is expressed in *E. coli*, e.g., strain PEB199. Expression of the GST-TOLL 25 fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity 45 chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

30

50

5

**EXAMPLE 3: EXPRESSION OF RECOMBINANT TOLL
PROTEIN IN COS CELLS**

10

To express the TOLL gene in COS cells, the pcDNA/Amp vector by Invitrogen

15

Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire TOLL protein and an HA tag (Wilson et al. (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

20

To construct the plasmid, the TOLL DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the TOLL coding sequence starting from the

25

initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the TOLL coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, MA). Preferably the two restriction sites chosen are different so that the TOLL gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5a, SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

30

COS cells are subsequently transfected with the TOLL-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsch, E. F., and Maniatis, T. 30 *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1989. The expression

40

45

50

55

5

of the TOLL polypeptide is detected by radiolabelling (^{35}S -methionine or ^{35}S -cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the

10 cells are labeled for 8 hours with ^{35}S -methionine (or ^{35}S -cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

10

15 Alternatively, DNA containing the TOLL coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the TOLL polypeptide is detected by radiolabelling and immunoprecipitation using a TOLL- specific monoclonal antibody.

20

25

Equivalents

30

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following

20 claims.

35

40

45

50

55

Claims

5

10

15

20

25

30

35

40

45

50

55

5

What is claimed:

1. An isolated nucleic acid molecule selected from the group consisting of:
 - (a) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:1;
 - (b) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:3; and
 - (c) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO: 4.
2. An isolated nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 2.
3. An isolated nucleic acid molecule comprising the nucleotide sequence contained in the plasmid deposited with ATCC® as Accession Number _____.
4. An isolated nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 2.

20

35

40

45

50

55

5

5. An isolated nucleic acid molecule selected from the group consisting of:

a) a nucleic acid molecule comprising a nucleotide sequence which is at least 60% homologous to the nucleotide sequence of SEQ ID NO:1, 3, or 4, or a complement thereof;

10

5 b) a nucleic acid molecule comprising a fragment of at least 491 nucleotides of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof;

15

c) a nucleic acid molecule which encodes a polypeptide comprising an amino acid sequence at least about 60% homologous to the amino acid sequence of

10 SEQ ID NO:2;

20

d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 15 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2; and

25

15 e) a nucleic acid molecule comprising a fragment of at least 15 nucleotides of a nucleic acid comprising the nucleotide sequence of SEQ ID NO: 4.

30

6. An isolated nucleic acid molecule which hybridizes to the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 under stringent conditions.

20

35

7. An isolated nucleic acid molecule comprising a nucleotide sequence which is complementary to the nucleotide sequence of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5.

40

25 8. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5, and a nucleotide sequence encoding a heterologous polypeptide.

45

9. A vector comprising the nucleic acid molecule of any one of claims 1, 2, 30 3, 4, or 5.

50

55

5

10. The vector of claim 9, which is an expression vector.

10

11. A host cell transfected with the expression vector of claim 10.

15

5 12. A method of producing a polypeptide comprising culturing the host cell
of claim 11 in an appropriate culture medium to, thereby, produce the polypeptide.

20

13. An isolated polypeptide selected from the group consisting of:
a) a fragment of a polypeptide comprising the amino acid sequence
10 of SEQ ID NO:2, wherein the fragment comprises at least 15 contiguous amino acids of
SEQ ID NO:2;
b) a naturally occurring allelic variant of a polypeptide comprising
the amino acid sequence of SEQ ID NO:2, wherein the polypeptide is encoded by a
25 nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of SEQ ID
NO:1, 3, or 4 under stringent conditions;
c) a polypeptide which is encoded by a nucleic acid molecule
30 comprising a nucleotide sequence which is at least 60% homologous to a nucleic acid
comprising the nucleotide sequence of SEQ ID NO:1, 3, or 4;
d) a polypeptide comprising an amino acid sequence which is at least
20 60% homologous to the amino acid sequence of SEQ ID NO:2.

35

14. The isolated polypeptide of claim 13 comprising the amino acid sequence
of SEQ ID NO:2.

40

25 15. The polypeptide of claim 13, further comprising heterologous amino acid
sequences.

50

45 16. An antibody which selectively binds to a polypeptide of claim 13.

55

5

17. A method for detecting the presence of a polypeptide of claim 13 in a sample comprising:

10

a) contacting the sample with a compound which selectively binds to the polypeptide; and

5

b) determining whether the compound binds to the polypeptide in the sample to thereby detect the presence of a polypeptide of claim 13 in the sample.

15

18. The method of claim 17, wherein the compound which binds to the polypeptide is an antibody.

20

19. A kit comprising a compound which selectively binds to a polypeptide of claim 13 and instructions for use.

25

20. A method for detecting the presence of a nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 in a sample comprising:

30

a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and

b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample to thereby detect the presence of a nucleic acid

20 molecule of any one of claims 1, 2, 3, 4, or 5 in the sample.

35

21. The method of claim 20, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

40

25 22. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 and instructions for use.

45

50

55

5

23. A method for identifying a compound which binds to a polypeptide of claim 13 comprising:

10

- a) contacting the polypeptide, or a cell expressing the polypeptide with a test compound; and
- 5 b) determining whether the polypeptide binds to the test compound.

15

24. The method of claim 23, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

20

- a) detection of binding by direct detection of test compound/polypeptide binding;
- b) detection of binding using a competition binding assay; and
- c) detection of binding using an assay for TOLL activity.

25

25. A method for modulating the activity of a polypeptide of claim 13 comprising contacting the polypeptide or a cell expressing the polypeptide with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

30

26. A method for identifying a compound which modulates the activity of a polypeptide of claim 13 comprising:

35

- a) contacting a polypeptide of claim 13 with a test compound; and
- b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

40

45

50

55

Human TOLL cDNA sequence:

GTCGACCACCGTCCGGACCGTGGGGTTGGATTTCAAAAGAGTAAA
 CCAGACCGTGACCAAGGTGAGACTAAGAAGTGGAGTCATGCTTCACACGG
 CCATATCATGCTGGCAGCCATTCTGGTCTGGCTGTGGTAAATCTTCATG
 GGATCCACCATTGGCTGCCCGCTCGCTGTGAGTGGCTCGCCAGAACAAATC
 TGTTAGCTGTACAGAAGGCAGTGTGATCGCCATCCCAGAGGGATTCCCATC
 GAAACCAAATCTTGGACCTCACTAAAAACAGGCTAAAAGCGTCAACCCCTG
 AAGAATTCATATCATATCCTCTGCTGGAAGAGATAGACTTGAGTGACAACAT
 CATTGCCAATGTGGAACCGAGGACATTCAAACTCTTAACTTGGCTTCCC
 TCCGCCAAAAGGCAATCGTCTAAAGCTGGCCCTTGGAGTATCACGGG
 GCTGTCCAATCTCACTAAGCTTGACATTAGTGAATAAGATTGTCAATTAC
 TAGACTACATGTTCCAAGATCTACATAACCTGAAGTCTAGAAGTGGGGGA
 CAATGATTGGTTATATATACACACAGGGCATTCACTGAGCTTACGTTGG
 AGCAGCTCACCCCTGGAGAAATGCAACTTAACAGCAGTACCAACAGAACCCCT
 CTCCCACTCCCGAGGCTCATCACCTGCATCTGAAGCATCTCAATATCAACA
 ATATGCCCTGTGATGCCCTTAAAGATTGTTCCACCTGAAACACCTAGAGATT
 GACTATTGGCCTTACTGGATATGATGCCCTGCCAATAGCCTCTACGGCTCAA
 CCTCACATCCCTTCAGTCACCAACACCAATCTGTCTACTGTACCCCTTGC
 CTTAACACCTGGTATAACCTGACTCACCTAACCTCTCTACAATCCCATCA
 GCACTATTGAAGCAGGCATGTTCTGACCTGATGCCCTCAGGAGCTTCAT
 ATAGTGGGGGCCAGCTTCGACCAATTGAGCCTCACTCCTCCAAGGGCTCCG
 CTTCCCTACGCGTGTCAATGTGTCTAGAACCTGCTGGAAACTTGGAGAGA
 ATGTCTCTCCTCCCTAGGGCTCTGGAGGTCTTGAGCATTAACAACACCCCT
 CTGGCCTGTGACTGCGCTTCTCTGGATCTTGAGCAGGACAGGCCACCCCTGCA
 GTTTGGGGCCAGCAACCTATGTGTGCTGGCCAGACACCATCCGTGAGAGG
 TCTTTCAAGGATTCATAGCACTGCCCTTCTTTTACTTTACCTGCAAAAAAA
 CCCAAATCCGTTGCCAGGATCAAGACAGCGGGATGTATGTTGCATCG
 CTAGCAATGCTGCTGGGAATGATACTTCACAGCCTCTAACGTGAAAGG
 ATTCGCTCAGATGTTTCTTATGCGAACAGGACCCCTATGTACATGACCG
 ACTCCAATGACACCAATTCCAATGGCACCAATGCCAATACTTTCCCTGGAC
 CTTAAAACAATACTGGGTCTACAGCTATGGCTGCTCACATTCCGGAGT
 GGTTTATTTGTTCTCTCCCTTGTGAGCCAGGGAAAGGCAAGC
 ACAAAACAGCATTGACCTTGAGTATGTCAGGGAAATGGCTGT
 TGTGGAAGGGGAGGTAGCTGGACCCAGGAGGTTCAACATGAAATGATTGA
 AGGCCCACCCCTCACATTACTGTCCTTGTCAATGTGGTAATCAGTAAGAC
 AGTATGGCACAGTAATTACTAGATTAAGAGGCAGCCATGTGCAGCTGCC
 TGTATCAAAGCAGGGCTATGGAGCAGGAGGACTTCCAATGGAGACTCTC
 CATCGAAAGGCAGGCAGGCAGGCATGTGTCAGAGCCCTCACACAGGGAT
 ACTAAGTGTGTCGCAATTCAGGACAATTTCAGCATTCTACCACTGCA
 TGAATATTGGCTCACACTCACGGACAATTTCAGCATTCTACCACTGCA
 AAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Fig. 1

SUBSTITUTE SHEET (RULE 26)

Human TOLL coding region:

ATGCTTCACACGGCCATATCATGCTGGCAGCCATTCCCTGGGTCTGGCTGTGGT
 GTTAATCTTCATGGGATCCACCAATTGGCTGCCCGCTCGTGTGAGTGCTCTG
 CCCAGAACAAATCTGTTAGCTGTACAGAAGGCAGATTGATGCCATCCCAGA
 GGGCATTCCCACATCGAAACCAAATCTTGGACCTCAGTAAAAACAGGCTAAA
 AGCGTCAACCCCTGAAGAACATTCAATCATATCTCTGCTGGAAGAGATAGACT
 TGAGTGACAACATCATTGCCAATGTGGAACAGGAGCATTCAACAATCTT
 TAACCTGCGTTCCCTCCGGCTAAAGGCAATGCTCTAAAGCTGGTCCCTTGG
 GAGTATTCACTGGGCTGTCCAATCTCACTAAGCTTGACATTAGTGAGAATAA
 GATTGTCATTTACTAGACTACATGTTCCAAGATCTACATAACCTGAAGTCTC
 TAGAAGTGGGGGACAATGATTGTTATATACACACAGGGCATTCACTGG
 GCTCTTAGCTTGGAGCAGCTCACCCTGGAGAAATGCAACTTAACAGCAGTA
 CCAACAGAACGGCCCTCTCCCACCTCCGAGCCTCATCAGCCTGCATCTGAAGC
 ATCTCAATATCAACAAATGCCCTGTGTATGCCCTTAAAAGATTGTTCCACCTG
 AAACACCTAGAGATTGACTATTGCCCTTACTGGATATGATGCCCTGCCAATAG
 CCTCTACGGTCTCAACCTCACATCCCTTCACTGACTCACCAACACCAATCTGTCTA
 CTGTACCCCTCCTTGCCTTAAACACCTGGTATACCTGACTCACCTAACCTCT
 CCTACAATCCCACACTATTGAAGCAGGCACTGGTCTGACCTGATCCGC
 CTTCAAGGGCTCCGCTTCACTGGTCTGCTCAATGTGTCTCAGAACCTGCTGG
 AAACATTGGAAGAGAAATGCTTCTCCCTCCCTAGGGCTCTGGAGGTCTTGAGC
 ATTAACAACAACCCCTGCCCTGTGACTGCCCTTCTCTGGATCTTGAGCG
 ACAGCCCACCCCTGCAGTTGGTGCCAGCAACCTATGTGTGCTGGCCAGAC
 ACCATCCGTGAGAGGCTTCAAGGATTCCATAGCACTGCCCTTCTTCTT
 CTTTACCTGCAAAAACCCAAAATCCGCTTGCCTCAGGATCAAGACAGCGGG
 ATGTATGTTGCATCGCTAGCAATGCTGCTGGAAATGATACCTCACAGCCTC
 CTTAACTGTGAAAGGATTGCTTCAGATCGTTTCTTATGCGAACAGGACCC
 CTATGTACATGACCGACTCCAATGACACCATTCCAATGGCACCAATGCCAAT
 ACTTTTCCCTGGACCTTAAACAAATACTGGTGTCTACAGCTATGGGCTGCTT
 CACATTCCGGAGTGGTTTATTTGTTCTCTCTTGTGTGGAGCCG
 AGGGAAAGGCAAGCACAACAGCATTGACCTTGAGTATGTGCCAGAAA
 AAACAATGGTGTGTTGTGGAGGGAGGTAGCTGGACCCAGGAGGTCAAC
 ATGAAAATGATTGA

Fig. 1 (continued)

Human TOLL protein sequence:

MLHTAISCWQPFLGLAVVLIFMGSTIGCPARCECSAQNKSVSCHRRRLIAIPEGIPI
ETKILDLSKNRLKSVNPEEFISYPPLLEEIDLSDNIIIANVEPGAFNNLFNRLSLRLKG
NRLKLVPLGVFTGLSNLTKLDISENKIVILLDYMFDQLHNLKSLEVGDNDLVYIS
HRAFSGLLSLEQLTLEKCNLTAVPTEALSHLRSLISLHLKHLNINNMPVYAFKRLF
HLKHLLEIDYWPLLDMMPPANSLYGLNLTSLSVTNTNLSTVPPFLAFKHLVYLTHLN
LSYNPISTIEAGMFSIDLRLQELHIVGAQLRTIEPHSFQGLRFLRVLNVSQNLLET
EENVFSSPRALEVLSINNNPLACDCRLLWILQRQPTLQFGQQQPMCAGPDTIRERS
FKDFHSTALSFYFTCKPKIRFAQDQDSGMVYCIASNAAGNDTFTASLTVKGFAS
DRFLYANRTPMYSNDTISNGTNANTPSLDLKTILVSTAMGCFTFLGVVLF
LLLFWWSRGKGKHKNSIDLEYVPRKNNGAVVEGEVAGPRRFNMKMI.

Fig. 1 (continued)

4/6

HMMER 2.1.1 (DEC 1998)

Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL)HMM file: /prod/ddm/seqanal/PPAM/pfam3.4/Pfam
Sequence file: /tmp/orfanal.28310.aa

Query: mill-toll

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
LRR	PF00560 Leucine Rich Repeat (2 copies)	183.5	3.3e-51	6

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
LRR	1/6	58	105	..	1	48 []	32.3 1.1e-05
LRR	2/6	106	153	..	1	48 []	43.9 3.7e-09
LRR	3/6	154	201	..	1	48 []	37.6 2.8e-07
LRR	4/6	202	250	..	1	48 []	17.1 0.42
LRR	5/6	251	297	..	1	48 []	34.9 1.8e-06
LRR	6/6	298	345	..	1	48 []	47.4 3.1e-10

Alignments of top-scoring domains:

LRR: domain 1 of 6, from 58 to 105: score 32.3, E = 1.1e-05

*->nLeeLdLsnNkLtslppgalsnLpnLeeLdLsnNnLtslppglfqnL

+ Ldls N+L+s+ p+ + + p Lee+dls+N+ ++pg+f nL

mill-toll 58 ETKILDLSKNRLKSVNPEEFISYPLEEEIDLSDNIIANVEPGAFN 104

k<-*

mill-toll 105 F 105

LRR: domain 2 of 6, from 106 to 153: score 43.9, E = 3.7e-09

*->nLeeLdLsnNkLtslppgalsnLpnLeeLdLsnNnLtslppglfqnL

nL++L+L +N+L+ +p g++ +L+nL++Ld+s N++ 1 ++ fq+L

mill-toll 106 NLRSERLKGNRKLKVPLGVFTGLSNLTKLIDISENKIVILLDYMFQDL 152

k<-*

mill-toll 153 H 153

LRR: domain 3 of 6, from 154 to 201: score 37.6, E = 2.8e-07

*->nLeeLdLsnNkLtslppgalsnLpnLeeLdLsnNnLtslppglfqnL

nL++L++ +N L ++ a+s+L +Le L L +nLt +p +++++L

mill-toll 154 NLKSLLEVGDNDLVYISHRAFGSGLLSLEQLTLEKCNLTAVPTEALSH 200

k<-*

+

mill-toll 201 R 201

Fig. 2

SUBSTITUTE SHEET (RULE 26)

LRR: domain 4 of 6, from 202 to 250: score 17.1, E = 0.42
 *->nLeeLdLsnNkLtslppgalsnLpnLeeLdLsnNnLt.slppg1fqn
 +L +L+L + ++ + p a++ L +L+L +++ L + p +++ +
 mill-toll 202 SLISLHLKHLNINNNMPVYAFKRLFHLKHLIDYWPPLLdMMMPANSLYG 248

Lk<-*
 L+
 mill-toll 249 LN 250

LRR: domain 5 of 6, from 251 to 297: score 34.9, E = 1.8e-06
 *->nLeeLdLsnNkLtslppgalsnLpnLeeLdLsnNnLt.slppg1fqnL
 L++L++n++L++p a++L L++L+Ls+N ++++ g f++L
 mill-toll 251 -LTSLSVTNTNLSTVPFLAFKHLVYLTHLNLSYNPISTIEAGMPSDL 296

k<-*
 mill-toll 297 I 297

LRR: domain 6 of 6, from 298 to 345: score 47.4, E = 3.1e-10
 *->nLeeLdLsnNkLtslppgalsnLpnLeeLdLsnNnLt.slppg1fqnL
 +L+L++ + +L+++p++++L+ L++L+s N L++l+++f++
 mill-toll 298 RLQELHIVGAQLRTIEPHSFQGLRFLRVLNVSQNLLETLEENVFSSP 344

k<-*
 +
 mill-toll 345 R 345

//

Fig. 2 (continued)

MOUSE TOLL DNA

ACTGTCTCTCTGTTACTGTTGGTCGTGAGTAAGACGCTGTAGAGTGACTCG
ATCACAAGGTTATCGGGCAGCTTGCGCAGCTGCCCTGTGTCAAAGCAGGG
TCCATGGAAAGCAGGAAGACTTCATGGAGACTGGCTGATTAGAGGCAGGCA
GGCATGTGTCAGAGCCCTCACACAGTGGGATACTAATTGTTGCATGCAAA
TATTGGCATTCTGGGATCTCAGCAATGAACCTGAAACCTTGGCTCATGCTGA
TGGACAATAATTCAACATTTCTACACTGCAAAACTAAAAGGAAAAAAAT
TAAAAGAACAAACCTACAGTGTAGGATTTACATATTAAAAGACACATTGCT
CTAAAACATACTCTACAGTCAAATTGTATTATTATCATTGTTAAAACCTT
GCATCATACAATAACTGTTGGTCAGCACCAAAAAGAGATCAATATATTCTTTT
TTTGAAACATATATGCTGTATATGTTAAAGCAATATGAATGAGAGGTTGT
GCTTTTAGTTACTCACCAGTATAGATCCAAGTGTGGTTCACCTCCCTTTACC
TGCAAGATAAACCTGACAATAGATCCCTGGAATACTAGGCAGAGATGTGTTGA
GATGTGTWTGTCGTAGGATGCCAAGAACAGAMCCAAGTCMAAACT
GCTCMACTCTGTTAACTCTGTTACTATAAAAGGCATGTGCCCTAGTTTG
ATACAAAAAAAAAARGGCCGC

Fig. 3